

Joint  
Transportation  
Research  
Program

**JTRP**

FHWA/IN/JHRP-96/18-01

Final Report

**AN INTEGRATED TRANSPORTATION  
LAND USE MODELING SYSTEM  
FOR INDIANA**

Part I

Andrew Ying-Ming Yen  
Jon D. Fricker

August 1998

Indiana  
Department  
of Transportation

Purdue  
University



Final Report

FHWA/IN/JHRP-96/18

AN INTEGRATED TRANSPORTATION LAND USE MODELING SYSTEM  
FOR INDIANA

PART I

by

Andrew Ying-Ming Yen  
Research Associate  
and  
Jon D. Fricker  
Professor of Civil Engineering

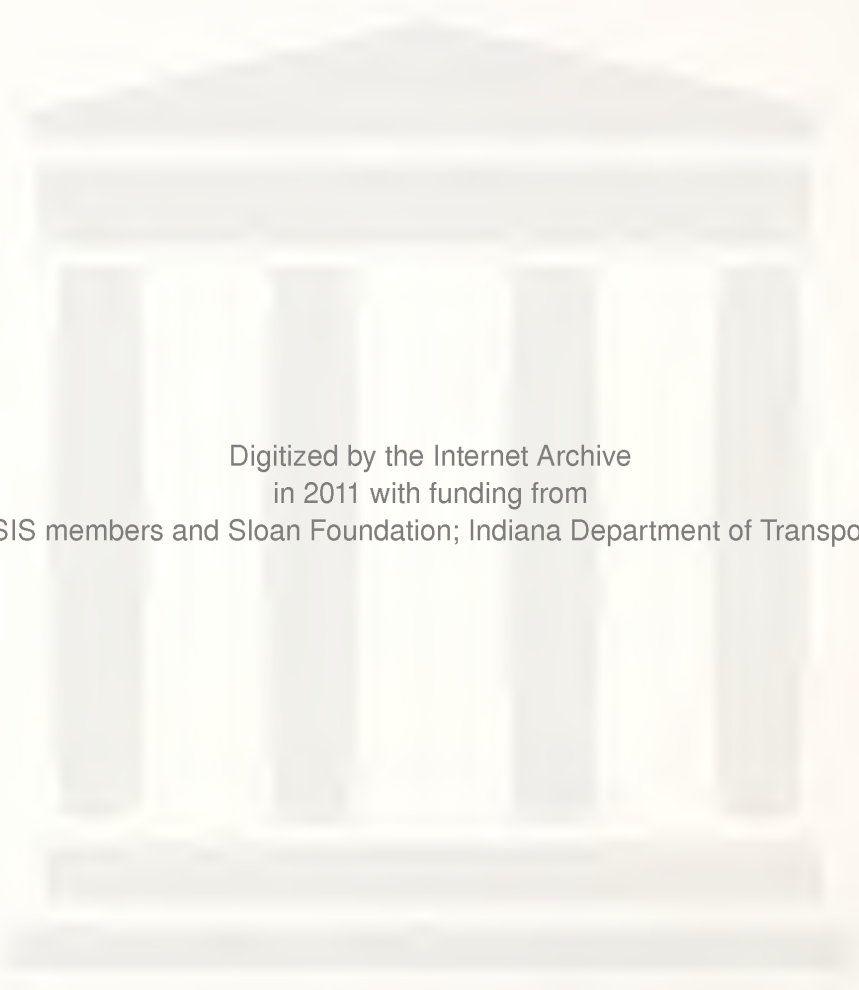
Purdue University  
School of Civil Engineering

JOINT TRANSPORTATION RESEARCH PROGRAM  
Project No. C-36-54ZZ  
File No. 3-3-52

Prepared in Cooperation with the  
Indiana Department of Transportation and the  
U.S. Department of Transportation  
Federal Highway Administration.

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Indiana Department of Transportation or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Purdue University  
West Lafayette IN 47907-1284  
August 1998



Digitized by the Internet Archive  
in 2011 with funding from  
LYRASIS members and Sloan Foundation; Indiana Department of Transportation

TECHNICAL REPORT STANDARD TITLE PAGE			
1. Report No. FHWA/IN/JHRP-96/18		2. Government Accession No.	
4. Title and Subtitle  An Integrated Transportation Land Use Modeling System for Indiana, Part I and Part II		3. Recipient's Catalog No.	
		5. Report Date August 1998	
7. Author(s)  Andrew Ying-Ming Yen and Jon D. Fricker		6. Performing Organization Code	
9. Performing Organization Name and Address Joint Transportation Research Program 1284 Civil Engineering Building Purdue University West Lafayette, Indiana 47907-1284		8. Performing Organization Report No.  FHWA/IN/JHRP-96/18	
12. Sponsoring Agency Name and Address  Indiana Department of Transportation State Office Building 100 North Senate Avenue Indianapolis, IN 46204		10. Work Unit No.	
		11. Contract or Grant No. HPR-2107	
		13. Type of Report and Period Covered  Final Report	
		14. Sponsoring Agency Code	
15. Supplementary Notes  Prepared in cooperation with the Indiana Department of Highways and Federal Highway Administration.			
16. Abstract  The principle objective of this research was to develop an integrated model to represent the interrelationships between land use and transportation, subject to the requirements of the ISTEA of 1991 and the CAAA of 1990. The integrated model includes two major parts: a land-use allocation module and a travel demand module. An interface module has also been built to transform data between these two modules. The land-use allocation module consists of a residential location model, an employment location model, a land use potential model, and a land consumption model. One unique feature of the residential and employment location models is that they simultaneously estimate passenger movement by work-to-home, home-to-shop and work-to-shop trips between zones before entering the travel demand module. Then the TRANPLAN-based travel demand module carries out trip generation and trip distribution stages for estimating home-based school, home-based other, non-home-based, external-internal and external-external trips. The land consumption model, which is significantly different from the LANCON procedure of DRAM/EMPAL, is based on micro-economic theory to simulate the profit-maximizing behavior of housing or landowners over time. The major purpose of the land consumption model is to satisfy the need to reach a balance between demand and supply in the housing market during each time period. The land consumption model is also influenced by Anas's work in CATLAS, which did not deal with employment location, but simultaneously considers housing location and employment location. The integrated model can be used to evaluate land use policies and transportation policies. Tests run for the Lafayette area have demonstrated that the model can be used to quantify positive and negative effects of long range transportation and land use plans.  The final report has two parts. Part I is the technical report that describes how the Indiana Transportation Land Use Modeling System (ITLUMS) was developed. Part II is a separate user's guide, which describes the FORTRAN programs that make up ITLUMS, the sequence in which the programs must be run, and the format of the input files needed by the ITLUMS programs.			
17. Key Words  land use, travel demand, integrated models, land use allocation, land use potentials.		18. Distribution Statement  No restrictions. This document is available to the public through the National Technical Information Service, Virginia, 22161	
19. Security Classif. (of this report)  Unclassified	20. Security Classif. (of this page)  Unclassified	21. No. of Pages Part I - 28 Part II - 158	22. Price



## TABLE OF CONTENTS

	Page
LIST OF TABLES .....	vi
LIST OF FIGURES .....	viii
CHAPTER 1. INTRODUCTION .....	1
CHAPTER 2. LITERATURE REVIEW .....	5
CHAPTER 3. MODEL STRUCTURE AND METHODOLOGY .....	16
3.1 Framework.....	16
3.2 Land-Use Allocation Module .....	18
3.3 Land Use Potential Study .....	26
3.3.1 Land Suitability Analysis .....	27
3.3.2 GIS Database .....	28
3.4 Land Consumption Model .....	28
3.5 Interface Module .....	38
3.6 Travel Demand Module .....	38
3.6.1 Simultaneous Location and Trip-distribution Procedures .....	38
3.6.2 Trip Generation .....	40
3.6.3 Trip Distribution .....	43
3.6.4 Modal Split .....	43
3.6.5 Traffic Assignment .....	46
3.7 Model Calibration .....	52
CHAPTER 4. DATA REQUIREMENTS AND ANALYSIS .....	66
4.1 Data Requirement .....	66

	Page
CHAPTER 5. CASE STUDY .....	76
5.1 Tippecanoe County, Indiana .....	76
5.1.1 Land Use Allocation Module .....	78
5.1.2 Land Use Potential Study .....	85
5.1.3 Land Consumption Model .....	87
5.1.4 Travel Demand Model .....	88
5.1.5 Policy Test 1 - Improvements of highway network for 2010 .....	91
5.1.6 Policy Test 2 - Changes in land use pattern .....	93
5.1.7 Policy Test 3 - Improvements of highway network for 2015 .....	94
CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS .....	119
6.1 Conclusions and Recommendations .....	119
6.2 Future Research .....	123
6.3 Implementation .....	124
LIST OF REFERENCES .....	125
Appendix A. Land-Use Potential Study – Scoresheet System Approach .....	132
Appendix B. Dempster-Shafer Evidence Theory .....	135



## LIST OF TABLES

Table	Page
3.1 Lot area (square feet) per dwelling unit, by district .....	60
3.2 Tippecanoe County, Indiana: 1994 data on auto occupancy .....	44
5.1 Summary of the Lafayette area in 1970, 1980, 1990,1993, 2000 .....	76
5.2 Summary of the Lafayette area by districts in 1980, 1985, 1990, 1993 .....	98
5.3 Calibrated parameters of the land use allocation module on 1980, 1990 and 1993 data .....	99
5.4 The average rents in Tippecanoe County townships in 1990 .....	100
5.5 Estimated coefficients for the existing housing supply submodel in the Lafayette area .....	100
5.6 Year 2010 population forecasts for Policy Test 1 .....	101
5.7 Differences in Year 2010 population forecasts for Policy Test 1 .....	101
5.8 Year 2010 Retail employment forecasts for Policy Test 1 .....	102
5.9 Differences in Year 2010 retail employment forecasts for Policy Test 1 .....	102
5.10 Comparisons of estimated average travel time based on Policy Test 1 and base case .....	103
5.11 Peak hour Volume/Capacity ratios on old US 231 in 1993 and in 2010 based on Policy Test 1 .....	103

Table	Page
5.12 Year 2010 Population forecasts Policy Test 2 .....	104
5.13 Year 2010 Retail employment forecasts for Policy Test 2 .....	104
5.14 Peak hour Volume/Capacity ratios on SR 26 in 1993 and in 2010 based on Policy Test 2 .....	105
5.15 Year 2015 Population forecasts for Policy Test 2 .....	105
5.16 Differences in population forecasts for Policy Test 3 .....	106
5.17 Year 2015 Retail employment forecasts for Policy Test 3 .....	106
5.18 Differences in Year 2015 retail employment forecasts for Policy Test 3 .....	107
5.19 Comparisons of estimated average travel time based on Policy Test 3 and base case .....	108
5.20 Peak hour Volume/Capacity ratios on Northwestern Avenue and US 52 in 1993 and in 2015 based on Policy Test 3 .....	108
A.1 Land use potential decision-making matrix .....	134
B.1 Some common evidential intervals .....	139
B.2 Probability measures for assessing land suitability (Dempster-Shafer Theory) .	145
B.3 Calculation of lower probability .....	149
B.4 Calculation of upper probability .....	149

## LIST OF FIGURES

Figure	Page
3.1 Framework of Land-Use/Transportation Model .....	62
3.2 Order of model operation for simulating the dynamics of interaction between Land use and Transportation .....	63
3.3 Structure of land use potential model .....	64
3.4 The framework of land consumption model .....	65
4.1 Forested Land, Tippecanoe County, IN .....	71
4.2 Distance to Sanitary Sewer System, Tippecanoe County, IN .....	72
4.3 Road System, Tippecanoe County, IN .....	73
4.4 Rail Lines & Airports, Tippecanoe County, IN .....	74
4.5 Current and Expected Land Use, Tippecanoe County, IN .....	75
5.1 Tippecanoe County, IN .....	109
5.2 District System, Tippecanoe County, IN .....	110
5.3 Order of model operation for simulating the dynamics of interaction between Land use and Transportation, Tippecanoe County, IN .....	77
5.4 Current and Expected Land Use, Tippecanoe County, IN .....	111
5.5 Point System Method, Tippecanoe County, IN .....	112

Figure	Page
5.6 Dempster-Shafer Theory, Tippecanoe County, IN .....	113
5.7 Policy Test 1, Tippecanoe County, IN .....	114
5.8 Changes in the road network by 2000, Tippecanoe County, IN .....	115
5.9 Policy Test 2, Tippecanoe County, IN .....	116
5.10 Policy Test 3, Tippecanoe County, IN .....	117
5.11 Changes in the road network by 2010, Tippecanoe County, IN .....	118
B.1 The probability assignment based on A evidence .....	140
B.2 The probability assignment based on B evidence .....	140
B.3 The probability assignment based on A and B evidences .....	140

## CHAPTER 1 INTRODUCTION

Recent legislation in the form of the Clean Air Act Amendments (CAAA) of 1990 and the Intermodal Surface Transportation Efficiency Act (ISTEA) of 1991 has placed increased emphasis on the use of travel demand simulation models in environmental and planning analysis. The CAAA demands that air quality forecasting must consider the interaction between land use and transportation and requires regional land use policies for mitigating congestion and environment impacts. The ISTEA requires long range transportation plans to quantify impacts on land use. For example, MPOs in non-attainment areas need to indicate whether new transport projects will have negative impacts on air quality (Wegener, 1994). The conventional travel demand simulation processes have, however, received strong criticism from environmental groups (NARC, 1992). These groups believe that current modeling practice overstates the air quality emissions benefits of highway improvements by showing speed increases without the additional induced travel that may result from improved highway infrastructure (Hawthorn, 1991). This induced travel would be generated from trips previously foregone, relocation of activities over a wider area, or new growth within the urban area responding to the improved transportation system performance. (Neuburger, 1971; Putman, 1983,1991; Orski, 1989; Wachs, 1989; Mackett, 1993)

The CAAA of 1990 and the ISTEA of 1991 unquestionably will affect ongoing transportation programs and practices. The CAAA of 1990 will have a major impact on transportation planning and project development in areas not meeting the National Ambient Air Quality Standards for ozone and carbon monoxide. The federal-aid highway programs in those areas will be revised as a result of the new law's provisions for highway sanctions, conformity, and transportation/air quality issues (Hawthorn, 1991). The ISTEA of 1991 also will place a greater emphasis on the use of travel demand simulation models in environmental and planning analysis. These two acts will have a profound effect on the future of land use and transportation interactions.

One of the most visible elements of public investment in recent decades has been the enormous expenditures on new highway construction. This same period has also witnessed large-scale shifts in population, both from rural areas to urban areas, and within the urban areas from city center to automobile-oriented suburban development. Although it has proven difficult to establish clear, complete, and explicit quantitative definitions of the causal relationships between these developments, it is clear that these relationships exist. (Putman, 1991; Mackett, 1993; Wegener, 1994; Giuliano, 1995; Waddell, 1995)

Land use covers a variety of activities, including residing, working and shopping; physical infrastructure such as workplaces and homes; and the outcomes of market processes, such as property and land values. All these activities can be influenced by changes in transport. Urban theory also shows us that people locate their houses and workplaces by trading off housing and commute costs and other factors (e.g., amenities, schools, etc.). Commuters choose their residential locations by considering both housing needs and workplace access. Likewise, employers choose work sites that are accessible to employees with tolerable time and monetary costs (Giuliano, 1995). Therefore, good transportation planning requires a good forecast of the type, magnitude, and location of urban activities (land use). Most often, this is accomplished by predicting how land use patterns will change with respect to base year conditions. Typically, the analyst will predict future population and employment data for the study area, then allocate these area totals to subareas, such as traffic analysis zones. This allocation is not only considered by the area-wide control totals, but it must also conform to zoning restrictions on type of land use in each subarea and reflect the characteristics associated with each subarea that affect land use location decisions.

The effects of a change in the level of activity -- for example, the number of residents living in an area on the number of trips made and the impact of a significant new or improved transport link on the level of activity -- are observable phenomena. The influence of changes in land use pattern on travel pattern tends to be easier to observe than the converse relationship, because

(a) travel patterns are affected primarily by land use patterns, while

(b) land use choices are influenced by many factors, only one of which is transportation. The effects of land use on transport can be modeled at both aggregate and disaggregate levels using well-established travel demand modeling techniques, such as trip generation, trip distribution, modal split, and traffic assignment. There is, however, no common view and generally accepted way to model the converse relationship. (Putman, 1983; Mackett, 1993)

To address the above issues, the principal objective of this research is to develop an improved integrated land-use/transportation model to represent the two-way relationships between land use and transportation, subject to the following constraints:

1. The requirements of the ISTEA of 1991 and the CAAA of 1990.
2. The acquisition of a statewide license by Indiana Department of Transportation (INDOT) for the travel demand modeling software TRANPLAN.
3. Metropolitan Planning Organizations (MPOs) and other potential users of the integrated model will not have (or could not reasonably be expected to collect) land use and economic data in a form or in the level of detail that land use models previously-developed need as input.

Therefore, it may be necessary to create a feedback mechanism to account for new travel patterns resulting from changes in the development pattern and the converse. This mechanism, where a land use allocation step is integrated with the usual four-step (trip generation, trip distribution, modal split and traffic assignment) travel demand model, is discussed in more detail in Chapter 3, "Model Structure and Methodology".

Chapter 2 contains the literature review for this research project and identifies the strengths and weaknesses of the current integrated Land-use/Transportation models.

Chapter 3 describes the overall structure of the integrated Land-use/Transportation model developed and the methodologies adopted in this research.

Chapter 4 describes data requirements and the process of analytical quantitative analysis.

Chapter 5 is a case study from Tippecanoe County, Indiana.

Chapter 6 contains the conclusions and recommendations from this research, as well as suggestions for further research.



## CHAPTER 2 LITERATURE REVIEW

Conventional four-step travel demand models, whether aggregate or disaggregate, cannot adequately represent the long-term impacts of changes in transport supply or land-use/transport policies (Putman, 1983; Mackett, 1993). In land use planning, a common approach is the Delphi process, which relies heavily on local planners or experts. This approach does not explicitly identify the interaction between land use and transportation and does not systematically evaluate the impact of transportation on land use or the converse, because it is based on judgment rather than analytical quantitative approaches. The application of land-use/transportation models has evolved over two decades. The ISTE and the CAAA have accelerated this evolution. In general, the integrated land-use/transportation models can be grouped into two types, aggregate and disaggregate. Wegener (1994) mentioned that land-use/transportation models must be operational, meaning that they have been calibrated, implemented and used for policy analysis for at least one metropolitan region. This excludes theoretical models presented only as a set of equations. The major operational land-use/transportation models in use today can be classified into the following groups, using a modification of Waddell's classification system (Waddell, 1995). (For a more detailed review of operational models, see Waddell, 1995 and Wegener, 1994.)

### 1. Models based on Lowry-type model

ITLUP, DRAM/EMPAL (Putman, 1983,1991)

HLFM II + (Horowitz, 1993)

LILT (Mackett, 1983)

POLIS (Prastacos, 1986)

### 2. Models based on Input/Output economic model

MEPLAN (Echenique et al., 1990)

TRANUS (de la Barra, 1989)

KIM (Kim, 1983,1990)

3. Models based on Random Utility/Discrete Choice theory

CATLAS (Anas, 1985)

5-LUT (Martinez, 1992)

Hamerslag (1994)

Waddell (1995)

4. Models using microsimulation techniques

NBER-HUDS (Kain, 1985)

MASTER (Mackett, 1990)

IRPUD (Wegener, 1986; Wegener et al., 1991)

5. Models integrated with GIS.

CUFM (Landis, 1993)

Batty (1992)

### **Models based on Lowry-type model**

The Lowry-Garin model (Lowry, 1964; Garin, 1966), a spatial interaction approach, has been applied to study a number of policy issues relevant to fast-growing regions. These include new transportation facilities, increased labor-force participation, and changes in zoning regulations. However, the analysis of transportation is hampered by the inability of the model to account for endogenous congestion and route choice, or to provide a basis for measuring the value of benefits produced (Berechman, 1988; Putman, 1983,1995). Furthermore, the spatial interaction approach lacked any appropriate grounding in terms of economic theories of activity location (Wegener, 1994; Waddell, 1995). Hence, Cochrane (1975) developed a “location surplus” notion to provide a satisfactory theoretical underpinning for economic theories of activity location in a Lowry-Garin type model. The derivation indicates that the trips that provide the trip-maker with the greatest net benefit are the ones chosen. Furthermore, Anas (1983) demonstrated that the multinomial logit model (MNL) derived from random utility maximization is, at an

equal level of aggregation, equivalent to the entropy and gravity models proposed by Wilson (1967, 1970).

Although the basic Lowry model is decades old and is not accepted in some research circles, variations on it continue to be used in newly-developed software packages. A model with widespread application in the U.S. used to distribute residential population and employment is called ITLUP (Integrated Transportation Land-Use Package) (Harvey and Deakin, 1993; Putman, 1974 and 1983). ITLUP combines two separate components: a land-use model and a transportation network model. In a later project, the land use model was revised by improving calibration techniques and by modifying the spatial-allocation formulae (Putman, 1991 and 1996). The land use model component includes two major submodels: EMPAL (EMPloyment ALlocation model) and DRAM (Disaggregated Residential Allocation Model). Beginning with regional trends, transportation facility descriptions, and data on the current location of employment and population, the EMPAL model forecasts the future location of employment of several types. Taking regional trends, transportation facility descriptions, and data on the current location of population, along with EMPAL's forecast of the future location of employment, the DRAM model forecasts the future location of households. ITLUP does not realistically reflect the role of household, business, developers, and government in making decisions about their respective choice in development process: location and investment (Waddell, 1994). However, ITLUP is a definite improvement because it permits consistent treatment of traffic congestion and is based on random utility. In order to overcome ITLUP's lack of the land market mechanism, Putman (1995) incorporates a procedure (LANCON) to calculate land consumption. EMPAL is modified to calculate location by employers, followed by DRAM's calculation of location demand by household. LANCON takes both these calculated demands and estimates the actual change in the amount of land, by zone, that will be used by each demand category. After the calculations are done for each demand category, the sum of land used is adjusted, using an increase/decrease in densities, to match the land available for such uses. The new release of ITLUP, called METROPILUS, is one of the few successful urban models (Wegener, 1994). Some limitations of Lowry-type models, including DRAM/EMPAL, are:

- 1) Workplace and housing location are assumed to be sequentially determined. Actually, they may be jointly determined (Waddell, 1995).
- 2) The models only consider the demand side of residential and employment location choices. They ignore (or make a crude representation of) the supply process of housing and workplace choices.
- 3) The models do not simulate the dynamic process (e.g., the bidding action for house or building) of changes in land prices or housing rents.
- 4) The models do not realistically reflect the role of households, businesses, developers, or governments in making decisions concerning the development process, for example, location choices and investment (Waddell, 1995).

The data required for DRAM/EMPAL consist of the following items:

- 1) 3-5 household categories, defined in terms of income (e.g., high, middle, and low income groups) by zonal level, whose parameters are individually estimated.
- 2) Employment is divided into 4-8 categories (for example, retail, manufacturing and other employment) by zonal level, whose parameters are individually estimated.
- 3) Total land area for each zone in the study area.
- 4) Vacant developable land for each zone in the study area
- 5) The percentage of developable land already developed for each zone.
- 6) Ratio of the amount of household type  $n$  ( $n = 1, 2, \dots, 5$ ) who are employees in employment type  $k$  to the total number of employees in employment type  $k$  ( $k = 1, 2, \dots, 8$ ) over the study area.
- 7) Ratio of the number of employees in employment type  $k$  ( $k = 1, 2, \dots, 8$ ) to the total number of households over the study area.
- 8) Amount of land that can be used for residential development by zonal level.
- 9) Amount of land that can be used for employment type  $k$  ( $k = 1, 2, \dots, 8$ ) development by zonal level.
- 10) Impedance (e.g., travel time or travel cost) between each O-D zone pair.
- 11) A network consists of streets and intersections in the study area.

Another Lowry-Garin type model is called HLFM II+ (Highway Land-Use Forecasting Model II+) (Horowitz, 1993). HLFM II+ helps identify developmental impacts in an urban area owing to improvements in the highway and transit systems. HLFM II+ extends the Lowry-Garin land-use model by permitting constraints on the total amount of land that may be allocated to various activities, by allowing for the possibility for agglomeration in services, and by representing constraints on population and service employment in any district. The traffic forecasting procedures in HLFM II+ are identical to those of the Quick Response System II for Windows (QRS II). HLFM II+ can find a combined Traffic/Land Use equilibrium (or stabilized ) solution. The stabilized condition means that there are no variations in activity locations and transportation costs within the same time period (e.g. five years, ten years or more). However, HLFM II+ also has the disadvantages of other Lowry-derivative models noted earlier. A more complete explanation of the stabilized situation between land use and transportation will be described in Chapter 3.

The required data for HLFM II+ consist of the following items:

- 1) One household category in the study area. Employment is divided into 2 categories ,basic and service (or retail) employment, by zonal level.
- 2) Ratio of population to the total number of employees over the study area.
- 3) Ratio of the number of service employees to the total population over the study area.
- 4) Ratio of the number of service employees to the total number of employees over the study area.
- 5) Amount of land that can be used for residential development by zonal level.
- 6) Amount of land that can be used for service development by zonal level.
- 7) Land area required for each service employee.
- 8) Persons per dwelling unit.
- 9) Value of time.
- 10) Impedance (e.g., travel time or travel cost) between each O-D zone pair.
- 11) A network consists of streets and intersections in the study area.

The LILT (Leeds Integrated Land-use Transport) model (Mackett, 1983) was originally developed for application to the city of Leeds. It combines the trip distribution and mode split stages of conventional travel demand model with a land-use model based on Lowry-type theory to locate housing and economic activity. The model was disaggregated by three modes of transport (car, public transport and walk), three socioeconomic population groups, and twelve industrial sectors (manufacturing, transport and communications, construction and so on). LILT uses processes similar to ITLUP to allocate population and employment (or economic activity) and to calculate the journey-to-work and journey-to-shop trips between zones. Trips for other purposes are calculated after the distributions of population and economic activity have been found, by applying trip-generation and trip-attraction rates based on the interzonal cost. The LILT model is an aggregate land-use/transportation model. It cannot explicitly represent the links between decision processes of individuals (for example, choice of transport mode, home location) over time, as it represents population and employment associated with a set of cross-section zonal data, without identifying the various changes to the individuals' decisions over time (Mackett, 1990).

The Projective Optimization Land Use Information System (POLIS) of the San Francisco Region, developed by Prastacos for the Association of Bay Area Governments, is a mathematical programming formulation of a Lowry-type model based on random utility and the given location of basic employment (Prastacos, 1986). POLIS shares the same limitations as other Lowry-type models.

### **Models based on Input/Output economic model**

Two well-known models based on Input/Output economic analysis are MEPLAN (Echenique et al., 1990; Hunt and Simmonds, 1993), developed by Marcial Echenique and Partners, and TRANUS (de la Barra, 1989), developed by Tomas de la Barra. In land use modeling, these two models use Input/Output economic techniques to capture the structure and linkages of economic activities. They reflect the operation of the land market and capture the effect of transportation improvements on land prices, land use and



development more realistically than Lowry-derivative models, which lack a land market mechanism. Once the equilibrium stage of a land use model is reached, MEPLAN transforms the resulting flows of labor, service and goods into physical flows of person trips (to work, to shop, or to school) and goods movements. MEPLAN uses a conventional transport model to do the model split by a hierarchical multinomial logit model and the traffic assignment by a multi-path probabilistic approach under capacity restraint. The transport model is iterated until convergence, that is, until there is no change in transport costs from one iteration to the next. The generating costs (disutilities) of the transport model are fed back into the land use model, thus influencing the location of activities in the next time period. MEPLAN is the second successful urban model and has been applied in more than a dozen countries in the world, including Greater London (Wegener, 1994).

Furthermore, Kim (Kim, 1983 and 1990; Rho and Kim, 1989) proposed a combined transportation-land use model incorporating goods movements in which zonal travel demand is endogenously determined together with link congestion costs, optimal amounts of production and resulting efficient densities of land uses, once the transportation network is given. The model is integrated into a consistent mathematical programming framework based on input/output analysis, random utility and bid-rent theory. It is worth noting that only the KIM model determines a general equilibrium between transportation and location based on the strict economic sense of equilibrium between supply and demand with endogenous prices (Wegener, 1994). KIM's approach, however, does not model physical stock, such as workplace (non-residential building such as factories, shopping centers and so forth) and housing, and land use development.

### **Models based on Random Utility/Discrete Choice theory**

The attractiveness of a discrete choice framework based on random utility maximization has increased since the contributions of McFadden (1973) and Domencich and McFadden (1975). Analysis of mode choice can be done using the standard techniques of multinomial probit, multinomial logit and nested logit models (Ben-Akiva and Lerman,

1985). Logit and nested logit approaches have also been applied to the choice of residential location (Martinez, 1992a and 1992b), the joint choice of travel mode and residential location by Lerman (1977), McFadden (1978), and Anas (1981, 1982, 1985), and the joint choice of residential location, workplace and housing tenure by Waddell (1995).

The Chicago Area Transportation/Land Use Analysis System (CATLAS) was developed by Anas (Anas and Duann, 1985) and was applied to Chicago. CATLAS is an economic urban simulation model that contains a simultaneous equation system to estimate housing and land values and achieve equilibrium. It is designed to evaluate the impact of changes in transportation policies on housing and land prices, on residential location and on mode choice patterns. CATLAS can be viewed as a hybrid of the land rent and land use models (following Alonso, (1964)) with the travel and location choice model (following McFadden, (1973)). It is an elegant approach, but it has some implementation difficulties that may limit its potential development for short-term and long-term forecasting. Some of its limitations are described below.

- 1) The assumption that workers choose their residential location based on a prior choice of workplace is unrealistic. Actually, Waddell (1995) and Giuliano (1995) mentioned that housing location and workplace may be jointly determined.
- 2) The model does not deal with employment location. Hence, the validity of the predictions of land price impacts on changes in land use or transportation strategies will be questionable. In general, households and businesses compete for location, and businesses often outbid households for location where they compete (Waddell, 1995).
- 3) The mathematical formulation of the model is exceedingly complex, requiring that a system of hundreds of simultaneous equations be solved for equilibrium. The approach creates a large computational burden.
- 4) The model is estimated in aggregate form because of the simultaneous equations in the implementation. Further disaggregation would probably substitute a microsimulation approach for the approach adopted within the CATLAS. The microsimulation procedure would require the development of a different model (Waddell, 1995).



The 5-LUT (5-Stage Land-Use Transport) model is developed by Martinez for Santiago de Chile (Martinez, 1992a). Martinez extends the conventional four-step travel demand model with a residential location choice (logit form) as the fifth level. The model integrated the bid-rent theory and discrete-choice approach for residential location choice. It also has many of the same problems discussed with respect to CATLAS. But this model warrants further attention for potential extension and application.

A Dutch approach uses a general-equilibrium model that includes land use, employment and residential location, and highway and public transport networks (Hamerslag, 1993 and 1994). The model is based on micro-economic theory under budget and time constraints. The model identifies how the dynamics of land use development are influenced by the interaction between transport infrastructure and land use development, and provides forecasting and sensitivity testing of policies for changes in highway and public transport networks. This micro-economic approach offers the greatest opportunities for overcoming existing land-use/transportation model limitations and may warrant further research and investigation.

### **Models using microsimulation techniques**

Aggregate land-use/transportation models cannot explicitly explain the links between the decision processes (housing choice, mode choice, route choice and so forth) of individuals over time. For this reason, the application of disaggregate microsimulation approaches to land-use/transportation models becomes attractive, and a computing approach based on microsimulation using a large sample of households becomes more efficient than one based on aggregate data (Waddell, 1995).

The NBER (National Bureau of Economic Research) model (Ingram et al., 1972) and the HUDS (Harvard Urban Development Simulation) model (Kain and Apgar, 1985) are economic urban simulation models, which have been applied to policy questions concerning the housing market. The NBER-HUDS models were developed to evaluate the impacts of spatially housing improvement programs (Kain and Apgar, 1985). The NBER-HUDS models employ a disequilibrium framework to obtain housing choices and

location rents within 50 distinct housing markets using a linear programming approach. One major weakness of the NBER-HUDS models is the assignment of households to housing units following the disequilibrium process rather than the better-established market clearing procedures in CATLAS (Anas and Duann, 1985). The NBER-HUDS models, moreover, do not consider the interaction between transportation, land use and the housing markets

MASTER (Micro-Analytical Simulation of Transport, Employment and Residence) has been developed in the U.K. by Roger Mackett (1990). It was originally developed because models such as LILT could not satisfactorily represent the complexity of responses to changes in the transport system. Another microsimulation model was developed by Wegener at the Institute of Spatial Planning at the University of Dortmund (IRPUD) (Wegener, 1986; Wegener et al., 1991). The IRPUD model starts with aggregate data, but uses microsimulation approaches with endogenous sampling in its housing market submodel. The major shortcomings associated with the microsimulation techniques are: (Waddell, 1995)

- 1) Many assumptions about individuals' decision processes must be made. This limits its credibility.
- 2) The errors involved in the aggregation of microsimulation model behavior are not easily identified.
- 3) Disaggregation of key actors (e.g., households and businesses) produces a high computing cost, although this problem continues to diminish..

### **Models integrated with GIS**

Geographic Information Systems (GIS) are computer-based systems that are used to store and manipulate geographic information. This technology has been touted as an essential tool for the effective use of geographic information (Aronoff, 1989). Since the 1970s, GIS has evolved as a means of assembling and analyzing diverse spatial data (Star and Estes, 1990; Laurini and Thompson, 1994). It has greatly enhanced the planner's ability to build and maintain the spatial database for planning and analysis. GIS also improves the ability to visualize the land use and transportation planning environments at

the urban, regional, state, and national levels of government. Some new urban models have started to utilize GIS techniques. The California Urban Futures Model (CUFM) was developed at the Institute of Urban and Regional Development of the University of California at Berkeley as a disaggregate model of housing development, and is one of the first urban models to take advantage of GIS capabilities (Landis, 1993). The CUFM is not a general purpose land use forecasting model, but it demonstrates significant potential benefits from integrating GIS techniques into the land-use/transportation models (Waddell, 1995). Another model using GIS is one developed by Batty (1992). The data exploration, calibration, prediction, and the display of results from the urban model are entirely dealt with within GIS. Their efforts are exploratory and simplistic at this stage. The model, however, does not yet consider the interaction between land use and transportation.



## CHAPTER 3

### MODEL STRUCTURE AND METHODOLOGY

#### 3.1 Framework

Figure 3.1 shows the overall framework of an integrated Land-use/Transportation model developed in this research. It illustrates the relationships between land use forecasting and transportation system modeling. The integrated model simulates the activities of the markets for land use and transport. For each zone within a metropolitan area, the model estimates the equilibrium housing rents (or price of land) resulting from the demand created by the activities in the zone and the supply of land or housing units (or floorspace) at any one time. In a similar way, the model estimates the equilibrium generalized cost of transport, including travel time and costs, that reflects the interaction between the demand for travel and the supply of transport at any time. The transport cost (or accessibility) is fed back to the land use module in the next time period, thereby affecting the equilibrium rents of housing units or price of land and the location of activities.

The integrated model allows comparison between two runs of the integrated model, one of which is a do-nothing case against which a policy can be tested. The model can be used to evaluate land use policies such as new industrial or commercial construction and transportation policies such as the improvement of the road network. During each time period, the integrated model will be iterated until convergence, that is, the model reaches a stabilized situation between the land use allocation module and the travel demand module. A stabilized situation means that there is no significant change in the distribution of population and employees over the study area in the land use allocation module, and no significant change in transport costs (or travel time) in the travel demand module from one iteration to the next.

Figure 3.2 (Echenique et al., 1990) shows a dynamic process of interaction between land use and transportation. In general, the changes in the land use patterns will

immediately affect the distributions of travel patterns. Thus, the changes in economic activities (or trips by purpose) resulting from the changes in land use pattern will be used as inputs in the travel demand module in the same time period as for the land use allocation module. Changes in the travel network, however, will take a longer time to influence the spatial patterns of land use. Therefore, the transport disutility (or accessibility) produced in the travel demand module will be used in the land use allocation module during the next time period.

The integrated model includes two major parts: a land-use allocation module and a travel demand module. There are interface links (or modules) built to transform data between these two modules.

- 1) The land-use allocation module: It contains a residential location model, an employment location model, a land use potential model, and a land consumption model. The purpose of this module is to estimate the spatial pattern of households (or population), employment, and density of occupation and floorspace. It simultaneously estimates the passenger movement by work-to-home, home-to-shop and work-to-shop trips between zones before entering the travel demand module.
- 2) The travel demand module: This module contains trip generation and trip distribution stages for estimating home-based school, home-based other, non-home-based, external-internal and external-external trips. It splits the trip matrices into individual modes and then assigns the vehicle trips on the links of network being analyzed. It includes capacity restraint on these links to represent congestion on the routes and considers the behavior of individual stochastic route choice. Because a statewide license was acquired by the Indiana Department of Transportation (INDOT), this study used TRANPLAN as the basis for the travel demand module.
- 3) The interface module: This module converts the generalized costs produced from the travel demand module into the accessibility between zones for use in the land-use allocation module. It operates in a reverse way, transforming output from the land-use allocation module such as the flows of activities between zones into the physical flows

of person trips, e.g., the peak-hour trip matrices by purpose, used in the travel demand module.

### **3.2 Land-Use Allocation Module**

One major concern of the land-use allocation module is the analysis of spatial interaction used to forecast the likely direction of population growth, employment locations and the distribution of economic activities in a region. Any such location model should incorporate at least the following properties.

- (1) It should help predict the impact of a change in employment opportunities (e.g., a plant opening, expansion or closing) upon the redistribution of population.
- (2) It should reflect the changes in accessibility to transportation services.
- (3) It should recognize the importance of transportation costs and nontransportation attributes (e.g., housing needs and neighborhood surrounding) in an individual's choice of household site.
- (4) It should obey population (or housing) and employment capacity constraints that are either determined by fixing residential and employment densities or are affected by the natural topography of the region.

The spatial interaction structure of land-use allocation module in this research has its origins in entropy-maximizing theory (Wilson, 1967 and 1970). However, Wilson's work lacked an appropriate grounding in the economic theories of activity location (Putman, 1995). Consequently, a 'surplus' notion was developed by Cochrane (1975) to provide a satisfactory theoretical basis for economic theory of activity location in a Lowry-Garin gravity type model. The derivation of 'surplus' is explicitly discussed by Cochrane (1975). Furthermore, Anas (1983) asserts that the multinomial logit model (MNL) derived from random utility maximization is, at an equal level of aggregation, equivalent to the entropy and gravity models proposed by Wilson.



The model developed in this research uses a multivariate attractiveness function and a two-parameter travel deterrence function (i.e., Tanner's function,  $f(t_{ij}) = t_{ij}^{-\alpha} \exp(-\beta t_{ij})$ ), analogous to Putman's (1995) equation forms. A consistent constraint procedure is used to satisfy specific activity constraints, including population (or housing) capacity, employment capacity and employment strategies. The core of the land-use allocation module includes two components: the residential location model and the employment location model.

### The residential location model

Following the 'surplus' concept proposed by Cochrane, a "location surplus" notion was developed by Putman (1995). In order to derive the functional form of the residential location model, it is assumed that the number of zones is large, say 100 or more (Putman, 1995). It is assumed that the underlying probability distribution is approximately exponential in the upper tail (see Cochrane, 1975). The surplus is defined as the difference between the probabilistic utility  $u_{ij}$ , the gross benefit of taking the trip, and a deterministic trip cost  $c_{ij}$  of actually taking the trip. The cost is a "generalized cost", including direct travel payment and travel time cost and so on. Therefore, the surplus is given as  $s_{ij} = u_{ij} - c_{ij}$ . And, by substitution (see Cochrane, 1975), we can calculate the probability of the surplus 's', given all the prior assumptions.

$$\Phi_{ij}(s) = \exp[-h A_j e^{-\lambda(s-m+c_{ij})}] \quad (3.1)$$

where  $\Phi_{ij}(s)$  is the cumulative distribution of the "location surplus" accruing from the optimal trip between zone i and zone j, and where

$h, \lambda, m$ : constants.

$A_j$ : a measure of the attractiveness of zone j.

Continuing through the derivation, Cochrane obtained, given  $O_i$  trips originating from zone i, the expected number of trips from zone i to zone j:



$$T_{ij} = \frac{O_i A_j e^{-\lambda C_{ij}}}{\sum_j A_j e^{-\lambda C_{ij}}} \quad (3.2)$$

This is the conventional form of the singly constrained spatial interaction model (i.e. gravity model). Further, the total surplus for all trips actually made is: (Cochrane, 1975)

$$S_T = (1/\lambda) \sum_i O_i [0.577 + \ln(h e^{\lambda m} \sum_j A_j e^{-\lambda C_{ij}})] \quad (3.3)$$

Thus by a series of “rather innocuous assumptions” regarding a utility maximization for trip-making process, a robust economic underpinning can be developed for a spatial interaction model formulation (Putman, 1995). The computational form of the residential location model developed here is a modified singly-constrained spatial interaction model. The actual form differs from equation (3.2) in that the attractiveness variable  $A_j$  is replaced by a multivariate formulation with a Cobb-Douglas functional form (Henderson and Quandt, 1980), namely

$$A_j = X_{1i}^{\alpha_1} X_{2i}^{\alpha_2} \dots X_{ni}^{\alpha_n}$$

where  $X_{1i}, X_{2i}, \dots, X_{ni}$  are the attributes of zone  $j$  such as available land, the number of dwelling units, etc., and  $\alpha_1, \alpha_2, \dots, \alpha_n$  are empirically calibrated parameters.

The application in the residential location model is an aggregate form of a multinomial logit model of location choice. The equation systems, analogous to Putman’s equation systems (1995), currently in use are described as follows.

$$P_i^n = \sum_j Q_j^n B_j^n w_i^n t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij}) \quad (3.4)$$

where

$$Q_j^n = \sum_k \alpha_{k,n} E_j^k \quad (3.5)$$

$$\text{and } B_j^n = [\sum_i w_i^n t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij})]^{-1} \quad (3.6)$$

$$\text{and } w_i^n = X_{1i}^{\alpha_1} X_{2i}^{\alpha_2} \dots X_{ni}^{\alpha_n} \quad (3.7)$$

where

$P_i^n$  : the population living in household type  $n$  (e.g., single-family) in zone  $i$ .

$E_j^k$  : the number of employees in employment type  $k$  (e.g., retail) who work in zone  $j$ .

$\alpha_{k,n}$  : the ratio of population in household type  $n$  who are employees in employment type  $k$  to the total number of employees in employment type  $k$  over the study area. (For example, if there are 100 persons living in household type  $n$  who are also employees in employment type  $k$  in the study area and the total number of employees in employment type  $k$  over the study area is 200 at the same time, then  $\alpha_{k,n} = 100/200 = 0.5$ ).

$t_{ij}$  : the disutility (travel time or travel cost or both) of travel between zone  $i$  and  $j$ .

$W_i^n$  : the attractiveness of residential zone  $i$  for population living in household type  $n$ , and

$X_{1i}^{\alpha_1}, X_{2i}^{\alpha_2}, \dots, X_{ni}^{\alpha_n}$  are the attributes of zone  $i$ , such as vacant available land, the number of households or dwelling units, the average rent of housing units.

$\beta_1^n, \beta_2^n, \alpha_1, \alpha_2, \dots, \alpha_n$ : empirically calibrated parameters.

The term  $B_j^n W_i^n t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij})$  represents the average probability that a commuter employed at zone  $j$  will choose to live in zone  $i$ .

Putman (1995) mentioned that the location surplus is a measure of the aggregate benefit households receive from the attributes of a residential zone. Because the utility perceived by a household can only be measured on an ordinal scale (i.e., it is not possible to determine the monetary value of the utility), the location surplus measure should be interpreted as an index number. The larger the value of the location surplus, the more utility a household receives from its choice of residential location.

The location surplus used in the residential location model can be derived by using two approaches -- indirect utility approach and direct integration approach. Both methods produce the same location surplus measures and have the same assumption that households tend to maximize their utilities when choosing residential location. For the first approach, the residential location model is interpreted as a multinomial logit model and the location surplus can be derived by calculating aggregate indirect utility (Ben-Akiva and Lerman, 1985). As to the second approach, the location surplus is found by directly

integrating the travel demand function in the residential location model (Cochrane, 1975; Williams, 1976).

### 1) Indirect utility approach

The residential attractiveness in the residential location model may be defined as in equation (3.7). Therefore, the indirect utility can be defined as:

$$V_{i,j}^n = \ln[w_i^n t_{ij}^{\beta_1^n} \exp(-\beta_1^n t_{ij})] \quad (3.8)$$

where

$V_{i,j}^n$  : indirect utility received by a type n household that resides in zone i with a head-of-household employed in zone j.

$t_{ij}$  : the travel time between zone i and j.

$\beta_1^n, \beta_2^n$  : empirically calibrated parameters.

The indirect utility in equation (3.8) is only for a single household. In order to find aggregate location surplus, it is necessary to sum the values of the indirect utility for all type n households.

$$LS^n = \sum_j \sum_k (\alpha_{k,n} E_j^k) \left[ \sum_i w_i^n t_{ij}^{\beta_1^n} \exp(-\beta_1^n t_{ij}) \right] \quad (3.9)$$

where

$LS^n$  : the location surplus for type n households.

### 2) Direct integration approach

In the residential location model, the number of work-to-home trips is calculated by a function of  $t_{ij}$  as follows:

$$T_{i,j}^n(t_{ij}) = \sum_k E_j^k \frac{w_i^n t_{ij}^{\beta_1^n} \exp(-\beta_1^n t_{ij})}{\sum_i w_i^n t_{ij}^{\beta_1^n} \exp(-\beta_1^n t_{ij})} \quad (3.10)$$

Given this travel demand function, the location surplus for a single household can be calculated by using a straight-line integration path (Williams, 1976). The following integral will be evaluated.

$$-\int_{c_{ij}}^{\infty} T_{i,j}^m(t_{ij}) dt \quad (3.11)$$

The solution to this integral is:

$$\frac{1}{\lambda} \ln [W_i^n t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij})] \quad (3.12)$$

where  $\lambda$ : a scale parameter

Equation (3.12) is identical to the equation (3.8), except for the scale parameter. In the residential location model, the scale parameter of the multinomial logit model is assumed to be one. This assumption is valid because the multinomial logit formulation in the residential location model is a linear-in-parameters logit model. In order to obtain comparable location surplus measures for each household type, it is necessary to assume that  $\lambda$  is set to a fixed value. Because utility and location surplus are measured on an ordinal scale, we can, without loss of generality, assume that  $\lambda$  is equal to one for all household types when the direct integration approach is used to calculate the location surplus. Then, an aggregate location surplus derived by using direct integration method has the identical form to that of equation (3.9), which is produced from the indirect utility process.

### The employment location model

The computational form of the employment allocation model is also a modified singly-constrained spatial interaction model and is derived using a way similar to that for the residential allocation model.

$$E_j^k = \sum_i T_i A_i^k S_j^k t_{ij}^{\beta_3^n} \exp(-\beta_4^n t_{ij}) \quad (3.13)$$

where

$$T_i = \delta_k (\sum_n P_i^n + \sum_{m \neq k} E_i^m + E_i^k) \quad (3.14)$$

$$\text{and } A_i^k = [\sum_j S_j^k t_{ij}^{\beta_3^n} \exp(-\beta_4^n t_{ij})]^{-1} \quad (3.15)$$

$$\text{and } S_j^k = Y_{1j}^{\gamma_1} Y_{2j}^{\gamma_2} \dots Y_{3j}^{\gamma_3} \quad (3.16)$$

where

$T_i$  : the ratio  $\delta_k$  multiplies the total sum of population and employees in zone i.

$P_i^n$  : the population living in household type n (e.g., single-family) in zone i.

$E_j^k$  : the number of employees in employment type k (e.g., retail) working in zone j.

$S_j^k$  : the attractiveness of zone j for employment type k.

$t_{ij}$  : the disutility (travel time or travel cost or both) of travel between zone i and j.

$\delta_k$  : the ratio of employees in employment type k to the total sum of population and employees over the study area. (For example, there are 800 persons living in the study area and 200 employees working there at the same time, if there is 40 employees in employment type k in this area, then  $\delta_k = 40/(800+200) = 0.04$ .)

$Y_{1j}^{\gamma_1} Y_{2j}^{\gamma_2} \dots Y_{3j}^{\gamma_3}$  are attributes of zone j, such as retail employment, available land and density of employment.

$\beta_3^n, \beta_4^n, \gamma_1, \gamma_2, \dots, \gamma_k$ : empirically calibrated parameters.

The term  $A_i^k S_j^k t_{ij}^{\beta_3^n} \exp(-\beta_4^n t_{ij})$  represents the average probability that a person living in zone i and a commuter employed at zone i will take part in an economic activity k (e.g., shopping etc.) in zone j.

### Population and employment constraints

In order to obtain a solution that satisfies the population constraints, a balanced constraint procedure is included in the model. It proceeds as follows. Suppose  $P_i$  is the total capacity of population in zone i. Then it is necessary to find the  $T_{ij}$  (journey-to-work trips) that satisfy

$$\alpha_i \sum_j T_{ij} \leq P_i \quad (3.17)$$

where  $\alpha_i$  is the ratio of population to employees in residential zone i, and

$$P_i = \mu_i^p Z_i^p A_i \quad (3.18)$$

where  $Z_i^p$  is an average density coefficient for dwelling units in zone i ( $Z_i^p$  can be calculated based on the criteria listed in Table 3.1 (APC, 1992)) that converts available residential land  $A_i$  (which can be estimated based on the output from the land use potential study) into dwelling units and  $\mu_i^p$  is the ratio of population to dwelling units in zone i. The population constraint for each zone in the residential location model is implemented by a computer program. When the program is run, the procedure of satisfying population constraints relies upon defining two distinct sets of zones. Let  $S_1$  be the set of zones whose populations are fixed or constrained, and let  $S_2$  be the set of zones whose populations are unconstrained. Initially, all zones would belong to  $S_2$ . Suppose that, after one iteration, constraint (3.17) is violated for a number of zones. Then these zones are allocated their maximum population capacity and are subsequently transferred to the set  $S_1$ . The surplus population of these zones is reconverted to employment and then reallocated back to its workplace by using another spatial interaction process. The surplus population is then added to the next increment of population, whereas those zones that have reached their capacities are not allocated any further population. The process is then reiterated on a constrained and unconstrained zone basis until a final equilibrium solution that does not violate the capacity constraints is obtained.

Furthermore, a maximum value constraint for a particular population type in a specified zone is also included in the residential location model. This constraint is:

$$\sum_n P_i^n \leq P_i \quad (3.19)$$

where  $P_i^n$ : the maximum value for population of household type n in zone i.

Likewise, the employment constraints for each zone can be derived using a process similar to that for the population constraints. These constraints are:

$$E_j = \mu_j^e Z_j^e A_j \quad (3.20)$$

$$\sum_m E_j^m \leq E_j \quad (3.21)$$

where

$E_j$  : the total employment capacity in zone j.

$Z_j^e$  : an average density coefficient for floorspace in zone j.

$A_j$  : available employment land in zone j.

$\mu_j^e$  : the ratio of employment to floorspace for zone j

$E_j^m$  : the maximum value for employment type m in zone j.

### 3.3 Land Use Potential Study

Historically, those who are charged with the responsibility for planning the future use of land have depended heavily on individual's skills and abilities to apply generalized planning principles to specific locales. Such a land use planning has been more often an art than a science, and often the subject of great controversy in the community (APC, 1981). Furthermore, traditional land use plans have tended to be demand-oriented and directional. For example, more residential or commercial or industrial use areas are added at the far edge of similar existing development and the quantities of these areas are calculated to meet the needs of population projected to a future date. Hence, one of major purposes in the land use potential study is to develop an improved decision making approach that avoids mistakes made in traditional land use plans.

Land use potential is a rigorous way to determine the effects of natural and man-made factors in determining how land may best be used in the future. A land use potential study consists of the following activities:

- 1) Develop well-supported relationships between land characteristics (or factors) and suitable land uses.



- 2) Divide the study area into cells that offer adequate detail without being too numerous.
- 3) Assemble and portray all the information required for the procedure.
- 4) Apply the land suitability relationships to the assembled data and graphically portray the results of the land use potential study.

### **3.3.1 Land Suitability Analysis**

Land suitability analysis plays an important role in many planning activities, from site selection to large-scale public or commercial or industrial land use developments, to environmental impact assessment. Some land may be physically unsuited to a given use or intensity of use, or the land may be better suited for another use. An unsuitable land use plan may repeat or exacerbate past mistakes. For example, unsuitable land use can inadvertently create unnecessary problems of storm water management, water supply and sewage disposal, and can lead to the loss of nonrenewable resources of energy, prime farmland, natural landscaping and open space.

Therefore, it is necessary to build an appropriate land suitability analysis approach and apply it to assess land use potentials among competing land uses in a rational and consistent way. One approach to integrate spatial database management systems (DBMS) and a land suitability indexing method is shown in Figure 3.3. GIS in this system is viewed as a database management tool that manipulates and extracts information from a spatial database as one of the possible inputs to the central component of the land suitability indexing model.

In this research, two land suitability indexing procedures are examined. The eight factors (or evidence) of interest consist of soil productivity, soil limitation, tendency to flood, forestation, sanitary sewer availability, accessibility, railroads and airports, and current and expected use, which includes projects that have at least received regulatory approval. The first procedure is a simple land suitability indexing activity using a weighted attribute type analysis. This procedure provides an assignment of points based on the value or characteristic of each factor for the region of interest (Appendix A). A second



procedure uses Dempster-Shafer evidence theory to calculate the potential indices of land suitability (Chang, 1991). The Dempster-Shafer procedure is applied to the same factors as those used in the first procedure (Appendix B). The two land suitability indexing procedures compute indices that assess available land resources to determine the potential for supporting specific activities, such as agricultural, residential, commercial, industrial, and open space land uses. The results from land suitability analysis will become the input for computing developable area for specific activities (e.g., residential, commercial or other development). The developable area is viewed as an important component for estimating the residential and service attractiveness of land-use allocation module. The above two approaches can help planners encourage land use patterns that satisfy the need for urban development and the need to preserve nonrenewable resources such as energy, farmland and open space.

### **3.3.2 GIS Database**

The database of eight land suitability factors is built as a raster-based structure that is consistent with the grid system proposed in Tippecanoe County by APC (1981). The spatial data can be transformed or digitized into a GIS database by using commercial packages such as ARC/INFO or TransCAD. In this study, the data exploration, manipulation, and the display of results from the land use potential study are performed entirely within TransCAD.

### **3.4 Land Consumption Model**

In Section 3.2, the residential location model generates people's choices of residential location. This is a demand side process and it is assumed that people choose the most attractive residential locations available. On the supply side, the land consumption model simulates profit-maximizing behavior on the part of housing or land owners. Three decisions are simulated. First, for the owner of an existing housing unit, the decision is whether to withdraw the housing from the market and keep it vacant or whether to offer it on the market by selling or renting it. Second, for the owner of vacant land, the decision

is whether to build new housing on that land or whether to postpone the decision to next time period. Third, if the owner of an old house or building perceives that demolishing the building and selling the land is more profitable than continuing to rent it out, then the owner will follow the demolish-and-sell strategy. Otherwise, the decision will be postponed to the next time period. The above simulation process for making decisions on renting a house or letting it remain vacant, building new house or demolishing old ones was originally used in CATLAS (Anas and Duann, 1985). In the land consumption model, it is assumed that there is a  $t$ -year ( $t \geq 1$ ) lag. The amount of new housing constructed and old stock demolished in a given time period depend on decisions made in last time period.

The land consumption model consists of a number of equations to be solved simultaneously for each time period (e.g., 5 years). These equations, similar to Anas's work in CATLAS (Anas and Duann 1985), are written as follows (equations (3.22)-(3.30)).

$$\begin{aligned} & \sum_j \bar{Q}_j^{n,t} B_j^{n,t} w_i^{n,t} t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij}) \\ &= S_i^{n,t} [U_{i,r}^{n,t} (R_i^{n,t}, \bar{X}_i^{n,t}, \bar{\eta}_r) + U_{i,o}^{n,t} (R_i^{n,t}, \bar{X}_i^{n,t}, \bar{\eta}_o)] \end{aligned} \quad (3.22)$$

where

$$\bar{Q}_j^{n,t} = \sum_k \delta_{k,n}^t E_j^{k,t} \quad (3.23)$$

$$\text{and } B_j^{n,t} = [\sum_i w_i^{n,t} t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij})]^{-1} \quad (3.24)$$

$$\text{and } w_i^{n,t} = X_{1i}^{\alpha_1} X_{2i}^{\alpha_2} \dots X_{ni}^{\alpha_n} \quad (3.25)$$

where

$\bar{Q}_j^{n,t}$  : the number of housing units of household type  $n$  (e.g., single-family) in zone  $i$  offered to those employees working in zone  $j$  at time period  $t$ .

$S_i^{n,t}$  : the number of housing units of household type  $n$  in zone  $i$  at time period  $t$ .

$U_{i,r}^{n,t}$  : the probability that a housing unit of household type  $n$  in zone  $i$  at time period  $t$  will be offered for renter occupancy, given the ongoing average rent  $R_i^{n,t}$ .

$U_{i,o}^{n,t}$  : the probability that a housing unit of household type  $n$  in zone  $i$  at time period  $t$  will be offered for owner occupancy, given the ongoing average rent  $R_i^{n,t}$ .

$R_i^{n,t}$  : the average monthly rent of housing units of household type  $n$  in zone  $i$  during time period  $t$ .

- $\bar{X}_i^{n,t}$  : a vector of the zone's characteristics related to the supply side, such as the age of an existing housing unit.  
 $\bar{\eta}_r$  : a set of empirically calibrated parameters on the supply side for renter occupancy.  
 $\bar{\eta}_o$  : a set of empirically calibrated parameters on the supply side for owner occupancy.  
 $E_j^{k,t}$  : the number of employees in employment type k (e.g., retail) who work in zone j at time period t.  
 $\delta_{k,n}^t$  : the ratio of the number of households for household type n to employees in employment type k over the study area at time period t.  
 $t_{ij}$  : the disutility (travel time or travel cost or both) of travel between zone i and j.  
 $W_i^{n,t}$  : the attractiveness of residential zone i for population living in household type n at time period t, and

$X_{1i}^{\alpha 1}, X_{2i}^{\alpha 2}, \dots, X_{ni}^{\alpha n}$  are the attributes of zone i, such as vacant available land, the number of households or housing units, the average rent of housing units and so on.

$\beta_1^n, \beta_2^n, \alpha 1, \alpha 2, \dots, \alpha n$ : empirically calibrated parameters on the demand side.

$$S_i^{n,t} = S_i^{n,t-1} + C_i^{n,t-1} - D_i^{n,t-1} \quad (3.26)$$

$$C_i^{n,t-1} = (Z_i^{n,t-1} * L_i^{n,t-1}) * U_i^{n,c} (R_{i,s}^{n,t-1}, s = 1, \dots, m; \bar{X}_i^{n,t-1}, r, \bar{\lambda}) \quad (3.27)$$

$$D_i^{n,t-1} = O_i^{n,t-1} * U_i^{n,d} (R_{i,s}^{n,t-1}, s = a_i^{n,t-1}, \dots, m; \bar{X}_i^{n,t-1}, r, \bar{\gamma}) \quad (3.28)$$

$$L_i^{n,t-1} = L_i^{n,t-2} - C_i^{n,t-2} / Z_i^{n,t-2} + D_i^{n,t-2} / Z_i^{n,t-2} \quad (3.29)$$

$$O_i^{n,t-1} = O_i^{n,t-2} - D_i^{n,t-2} + A_i^{n,t-2} \quad (3.30)$$

where

$C_i^{n,t-1}$  : the expected number of new housing units constructed in zone i during time period t-1.

$D_i^{n,t-1}$  : the expected number of old housing units demolished in zone i during time period t-1.

$L_i^{n,t-1}$  : the area of vacant land available for household type n in zone i during time period t-1.

$Z_i^{n,t-1}$  : an average density coefficient for housing units in household type n in zone i during time period t-1. ( $Z_i^{n,t-1}$  can be calculated based on the criteria listed in Table 3.1)

$U_i^{n,c}$  : the expected proportion of potential dwellings that will be built during time period t-1.

$O_i^{n,t-1}$  : the number of old dwellings (e.g., over thirty years) for household type n eligible for demolition during time period t-1.

$U_i^{n,d}$  : the expected proportion of old dwellings to be demolished during time period t-1.

- $r$  : interest rate.  
 $\bar{\lambda}$  : a set of empirically calibrated parameters.  
 $\bar{\eta}$  : a set of empirically calibrated parameters.  
 $A_i^{n,j-2}$  : the number of dwellings of household type  $n$  in zone  $i$  aging into the over thirty years category.

The left side of equation (3.22) is similar to the right side of equation (3.4), where the term  $B_j^n w_i^n t_{ij}^{\beta_j^n} \exp(-\beta_i^n t_{ij})$  is the average probability of a commuter employed at zone  $j$  will choose to live in zone  $i$ . The equations in (3.22) are solved simultaneously for each time period  $t$  to obtain the values of the rent vectors  $\bar{R}_i^{n,t} = [R_1^{n,t}, R_2^{n,t}, \dots, R_i^{n,t}]$ , where  $R_i^{n,t}$  is the average monthly rent of the housing units in household type  $n$  in zone  $i$  during time period  $t$ .  $U_{i,r}^{n,t}(\cdot)$  and  $U_{i,o}^{n,t}(\cdot)$  are the expected proportions of the available dwellings  $S_i^{n,t}$  that will be offered for renter occupancy and owner occupancy during time period  $t$  respectively. And,  $1 - U_{i,r}^{n,t}(\cdot) - U_{i,o}^{n,t}(\cdot)$  is the probability to remain vacant. Equation (3.22) states that expected demand equals expected supply in each zone and in each time period  $t$ . Furthermore, equations (3.24) and (3.25) are similar to equations (3.6) and (3.7) in Section 3.2.

Equation (3.26) states that the number of dwellings in time period  $t-1$  increases by the expected number of new dwellings constructed,  $C_i^{n,t-1}$ , and decreases by the number of dwellings demolished,  $D_i^{n,t-1}$ , during the same time period.

Equation (3.27) represents the expected number of housing units to be built in a given time period  $t-1$ . The term  $(Z_i^{n,t-1} * L_i^{n,t-1})$  is the potential new housing that can be constructed in zone  $i$ . The function,  $U_i^{n,c}(\cdot)$ , is derived from the owner or developer's profitability decision. The decision depends on the stream of annual rent each dwelling is expected to earn over the dwelling's lifetime  $m$ , the supply side characteristics  $\bar{X}_i^{n,t-1}$ , the market interest rate  $r$ , and  $\bar{\eta}$ , a vector of parameters to be estimated.

Equation (3.28) estimates the expected number of demolitions during time period  $t-1$ .  $U_i^{n,d}(\cdot)$  is a function of the stream of annual rents that can be accrued from the average

old dwelling in household type  $n$  in zone  $i$  over its remaining lifetime, the supply side characteristics  $\bar{X}_i^{n,t-1}$ , the interest rate  $r$ , and  $\bar{\eta}$ , a vector of parameters to be estimated. The age of the average old dwelling in household type  $n$  in zone  $i$  is  $a_i^{n,t-1}$ .

Equation (3.29) adjusts the amount of vacant land available in a zone based on the processes of land taken up by new construction and land opened up by demolitions.

Equation (3.30) updates the number of dwellings eligible for demolition by adding  $A_i^{n,t-2}$ .  $A_i^{n,t-2}$  is calculated from a simple cohort-survival model for each household type  $n$  in zone  $i$ .

Finally, the land consumption model adjusts the rent of housing (or price of land) until the expected demand equals the expected supply (or the fixed supply) in each zone and at each time period  $t$ . That means that the demand and supply sides of a real estate market tend towards a “temporary equilibrium”. Because changes occurred in travel characteristics (e.g., a new transit system) or employment locations, etc., they will shift the study region to a new “temporary equilibrium” in the next time period (Anas and Duann, 1985).

The derivation of the left side of equation (3.22) has been described in Section 3.2. The right side of equation (3.22) will be derived in the submodel to adjust existing housing supply. The derivation of equations (3.27)–(3.28) is described in the housing stock adjustment submodel. The processes represented in the following submodels are analogous to Anas’s work in CATLAS (Anas and Duann, 1985). A framework of the land consumption model is shown in Figure 3.4.

### Submodel to adjust existing housing supply

In this submodel, the owner of an existing housing must decide whether to keep the housing vacant or offer it on the market for sale or rent, or occupy it himself at each time period. The decision is based on profitability. Suppose that an existing housing unit is offered for rent. Then the profit from this decision can be computed as follows.

$$\Delta_{i,1}^{n,t} = R_i^{n,t} - M_{i,1}^{n,t} + \mathcal{E}_{i,1}^{n,t} \quad (3.31)$$

If it remains vacant, then the financial loss is:

$$\Delta_{i,2}^{n,t} = -M_{i,2}^{n,t} + \mathcal{E}_{i,2}^{n,t} \quad (3.32)$$

where

$R_i^{n,t}$  : the average monthly rent for housing units of household type n (e.g., single-family) in zone i during time period t collected from census data.

$M_{i,1}^{n,t}$  : the average monthly cost of maintaining an existing housing unit of household type n if it is occupied during time period t.

$M_{i,2}^{n,t}$  : the average monthly cost of maintaining an existing housing unit of household type n if it is vacant during time period t.

$\mathcal{E}_{i,1}^{n,t}, \mathcal{E}_{i,2}^{n,t}$  : random error terms.

In general, the  $M_{i,1}^{n,t}$  will be higher than  $M_{i,2}^{n,t}$ , and the difference between them will depend on the type and location of the dwelling's neighborhood. The differential profit is:

$$\Delta_{i,1}^{n,t} - \Delta_{i,2}^{n,t} = R_i^{n,t} - (M_{i,1}^{n,t} - M_{i,2}^{n,t}) + \mathcal{E}_{i,1}^{n,t} - \mathcal{E}_{i,2}^{n,t} \quad (3.33)$$

The value of  $(M_{i,1}^{n,t} - M_{i,2}^{n,t})$  is, however, not directly available in the data. Because it depends on zone (or neighborhood) characteristics, such as the age and the location of housing units, it can be estimated by a function of these characteristics (Anas and Duann, 1985). We can assume

$$(M_{i,1}^{n,t} - M_{i,2}^{n,t}) = \sum_{m=1}^N \eta_m X_{i,m}^{n,t} \quad (3.34)$$

where

$X_{i,m}^{n,t}$  : the supply side zone characteristics of housing units of household type n in zone i during time period t, such as the location of the housing unit and the percentage of developed land.

$\eta_m$  's are the coefficients to be estimated.

Thus, the probabilities  $U_{i,r}^{n,t}$  and  $U_{i,o}^{n,t}$  that an existing housing unit will be offered for renter occupancy and owner occupancy can be derived using the following two approaches.

1) Linear regression method.

$$U_{i,r}^{n,t} = \eta_{r,0} R_i^{n,t} + \sum_{m=1}^N \eta_{r,m} X_{i,m}^{n,t} + \varepsilon_{i,r}^{n,t} \quad (3.35)$$

$$U_{i,o}^{n,t} = \eta_{o,0} R_i^{n,t} + \sum_{m=1}^N \eta_{o,m} X_{i,m}^{n,t} + \varepsilon_{i,o}^{n,t} \quad (3.36)$$

and satisfy

$$U_{i,r}^{n,t} + U_{i,o}^{n,t} + U_{i,v}^{n,t} = 1.0 \quad (3.37)$$

where

$U_{i,r}^{n,t}$  : the probability that a housing unit of household type n in zone i at time period t will be offered for renter occupancy, given the ongoing average rent  $R_i^{n,t}$ .

$U_{i,o}^{n,t}$  : the probability that a housing unit of household type n in zone i at time period t will be offered for owner occupancy, given the ongoing average rent  $R_i^{n,t}$ .

$U_{i,v}^{n,t}$  : the probability that a housing unit of household type n in zone i will remain vacant at time period t.

$\eta_{r,0}, \eta_{r,1}, \dots, \eta_{r,N}$  and  $\eta_{o,0}, \eta_{o,1}, \dots, \eta_{o,N}$  are empirically estimated coefficients.

$\varepsilon_{i,r}^{n,t}$  and  $\varepsilon_{i,o}^{n,t}$  are random error terms.

2) Logit model

$$U_{i,r}^{n,t} = \frac{U_1}{U_1 + U_2 + U_3} \quad (3.38)$$



$$U_{i,o}^{n,i} = \frac{U2}{U1 + U2 + U3} \quad (3.39)$$

$$U_{i,v}^{n,i} = \frac{U3}{U1 + U2 + U3} \quad (3.40)$$

where

$$\begin{aligned} U1 &= \exp(\eta_{r,0} R_i^{n,i} + \sum_{m=1}^N \eta_{r,m} X_{i,m}^{n,i}) \\ U2 &= \exp(\eta_{o,0} R_i^{n,i} + \sum_{m=1}^N \eta_{o,m} X_{i,m}^{n,i}) \\ U3 &= \exp(\eta_{v,0} R_i^{n,i} + \sum_{m=1}^N \eta_{v,m} X_{i,m}^{n,i}) \end{aligned}$$

Equations (3.38)-(3.40) comprise the multinomial logit model, where  $\eta_{r,0}, \eta_{r,1}, \dots, \eta_{r,N}, \eta_{o,0}, \eta_{o,1}, \dots, \eta_{o,N}$ , and  $\eta_{v,0}, \eta_{v,1}, \dots, \eta_{v,N}$  are empirically estimated coefficients.

### The housing stock adjustment submodel

The adjustments of housing stock are performed from one time period to the next. Only the construction of new dwellings on vacant land and the demolition of old dwellings are considered. Anas (Anas and Duann, 1985) mentioned that these decisions depend crucially on “present value of profit” (PVP) that can be calculated from a dwelling over its remaining lifetime. Suppose the average lifetime of a dwelling in zone  $i$  is  $m$  years and its average age is  $a_i$ . Then the present value of profit, (PVP) $_{i,j,a_i}$ , can be obtained from a rental decision now (time  $t$ ) until the end of a dwelling’s lifetime.

$$(PVP)_{i,j,a_i} = \sum_{s=a_i}^m \frac{(R'_{i,s} - M'_{i,1,s})(U'_{i,r} + U'_{i,o}) + (-M'_{i,2,s})(1 - U'_{i,r} - U'_{i,o})}{(1+r)^{s-a_i}} \quad (3.41)$$



If a dwelling can be constructed on vacant land in zone i, then the profit accrued from construction will be represented as follows.

$$\Delta_{i,c,t} = (PVP)_{i,t,1} + RS_{i,m}/(1+r)^m - C_{i,t} + \mathcal{E}_{i,t}^c \quad (3.42)$$

where

$RS_{i,m}$  : the resale value of a dwelling unit m years from now in zone i.

$C_{i,t}$  : the current construction cost for the dwelling in zone i.

$\mathcal{E}_{i,t}^c$  : a random error term

If the land remains vacant, the profit will be equal to the land price, which incorporates the present value of all future taxes and other expenses. The profit,  $\Delta_{i,v,t}$ , under this situation will be written as:

$$\Delta_{i,v,t} = V_{i,t} + \mathcal{E}_{i,t}^v \quad (3.43)$$

where  $V_{i,t}$  is the land price and  $\mathcal{E}_{i,t}^v$  is a random error term. The present value of profit,  $(PVP)_{i,t,1}$ , in equation (3.42) can be rewritten as:

$$(PVP)_{i,t,1} = \sum_{s=1}^m \frac{(R_{i,s}^t)(U_{i,r}^t + U_{i,o}^t)}{(1+r)^{s-1}} + \sum_{s=1}^m \frac{(M_{i,2,s}^t - M_{i,1,s}^t)(U_{i,r}^t + U_{i,o}^t) - M_{i,2,s}^t}{(1+r)^{s-1}} \quad (3.44)$$

The differential profit between the decision to construct a new dwelling versus keeping the land vacant can be described as:

$$\Delta_{i,c,t} - \Delta_{i,v,t} = \sum_{s=1}^m \frac{(R_{i,s}^t)(U_{i,r}^t + U_{i,o}^t)}{(1+r)^{s-1}} + \sum_{m=1}^N \lambda_m X_{i,m}^{n,t} + \mathcal{E}_{i,t}^c - \mathcal{E}_{i,t}^v \quad (3.45)$$

The term  $\sum_{m=1}^N \lambda_m X_{i,m}^{n,t}$  stands for the second summation in equation (3.44) plus  $V_{i,t}$ , but these data may not always be available. Thus a function of the supply side variables are used to represent these quantities.  $\lambda_0, \dots, \lambda_m$  are coefficients to be estimated. Under the above assumptions, the probability of a new dwelling that will be constructed on vacant land in zone i can be computed by a binary logit model.

$$U_i^{t,c} = \frac{\exp \left[ \lambda_0 \left[ \sum_{s=1}^m (R'_{i,s}) (U'_{i,r} + U'_{i,o}) / (1+r)^{s-1} \right] + \sum_{m=1}^N \lambda_m X_{i,m}^{n,t} \right]}{1 + \exp \left[ \lambda_0 \left[ \sum_{s=1}^m (R'_{i,s}) (U'_{i,r} + U'_{i,o}) / (1+r)^{s-1} \right] + \sum_{m=1}^N \lambda_m X_{i,m}^{n,t} \right]} \quad (3.46)$$

In a similar way, the probability of demolishing the average old dwelling in zone i can be derived by a binary logit model.

$$U_i^{t,d} = \frac{1}{1 + \exp \left[ \gamma_0 \sum_{s=a}^m (R'_{i,s}) (U'_{i,r} + U'_{i,o}) / (1+r)^{s-1} \right] + \sum_{m=1}^N \gamma_m X_{i,m}^{n,t} \right]} \quad (3.47)$$

where

$\gamma_0, \dots, \gamma_m$  are empirically estimated coefficients.

### Market clearing equilibrium during each time period

The equation (3.22), which is solved for the market clearing rent vector  $\bar{R}_i^{n,t} = [R_i^{n,t}, R_2^{n,t}, \dots, R_t^{n,t}]$ , is the core of land consumption model during each time period. Anas (1982) has demonstrated that the system of equations (equations (3.22)-(3.30)), given  $\bar{Q}_j^{n,t}, \delta_{k,n}^t, w_i^{n,t}, t_y, S_i^{n,t}, \bar{X}_i^{n,t}, \beta_1^n, \beta_2^n, \alpha_1, \alpha_2, \dots, \alpha_n$  and  $\bar{\gamma}$ , can be solved for a unique and stable equilibrium rent vector  $\bar{R}_i^{n,t}$  which clears the market during each time period.

### 3.5 Interface Module

This module is a bridge between the land use allocation module and the travel demand module. As discussed in Section 3.2, three types of trips -- home-based work (HBW), home-based shopping and work-based shopping -- are generated in the land-use allocation module. In the interface module, these three types of trips will be converted into HBW and home-based other (HBO) person trip matrices between each O-D pair. These two purpose's matrices have the same format as the home-based school (HBS), nonhome-based (NHB), external-internal (E-I), and external-external (E-E) matrices created in the trip-generation and trip-distribution stages of the TRANPLAN-based travel demand module. In general, the changes in the land use patterns will immediately influence the distributions of travel demand patterns. Thus, the HBW and HBO person trip matrices will be used as inputs in the travel demand module in the same time period as for the land use allocation module (see Figure 3.2).

In reverse, a major purpose of the travel demand module is to derive the transport disutilities, travel time or cost, between each zone pair. These transport disutilities are transformed into the disutilities of flows of activities used in the land use allocation module. In general, the changes in travel network will need a longer time to affect the spatial patterns of land use. Therefore, the transport disutilities produced in the travel demand module will be used in the land use allocation module at the next time period (see Figure 3.2).

### 3.6 Travel demand module

#### 3.6.1 Simultaneous Location and Trip-distribution Procedures

The operations in the land-use allocation module involve trips generated implicitly in the residential location model and the employment location model. The approaches can be

used as a substitute for the standard procedures for trip generation and distribution in conventional travel demand model. In Section 3.2, equation (3.4)

$$P_i^n = \sum_j Q_j^n B_j^n w_i^n t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij})$$

produces simultaneous forecasts of residence location and trip distribution for work-to-home trips, where the term  $B_j^n w_i^n t_{ij}^{\beta_2^n} \exp(-\beta_1^n t_{ij})$  is the probability of a work trip between zone i and zone j. Likewise, equation (3.13)

$$E_j^k = \sum_i T_i^k A_i^k S_j^k t_{ij}^{\beta_3^n} \exp(-\beta_4^n t_{ij})$$

generates simultaneous forecasts of the employment location and trip distribution for home-to-shop and work-to-shop trips. Therefore, in the land-use allocation module, three types of matrices (home-based work, home-based shopping and work-based shopping) are generated. Under the above simulation process, the residential location model calculates residential location demand by nonretail and retail (or basic and nonbasic) employees, followed the calculation of location demand by retail (or nonbasic) employment in the employment location model. The particular virtue of this joint calculation process is that the trip matrix is consistent with the spatial distribution of tripmakers, including their trip origins and trip destinations. This process differs significantly from the traditional procedure, which utilizes multiple regression or cross-classification in the trip generation stage, then uses a gravity model to produce a trip matrix.

By constraining a home-based work trip matrix at the work end of the trip, an attraction-constrained or destination-constrained procedure is adopted. For example, the procedure in the residential location model is an attraction-constrained process. In a similar way, a constraint on the home end of the trip means a production-constrained or origin-constrained procedure. The procedure in the employment location model that estimates home-based shop trips is a production-constrained process. Sometimes, it is necessary to constrain both ends of the trip. This is called a doubly constrained approach.

The home-based other (HBO), home-based school (HBS), nonhome-based (NHB), external-internal (E-I), and external-external (E-E) trips for trip-generation and trip-distribution stages will be dealt with using conventional travel demand approaches. These procedures are described in more detail in Sections 3.6.2 and 3.6.3.

### 3.6.2 Trip Generation

The trip rates method for trip productions and linear regression method for trip attractions described in TETUP (1994) are used to estimate HBO, NHB and E-I or I-E trips.

#### 1) Trip production

The cross-classification method described in TETUP (1994) is used to predict the number of trips produced in each zone. In general, the household trip production rates is stratified by household size and income, or by household size and auto ownership. Thus, the HBO trip productions and NHB trip productions in each zone can be computed using the following equations (3.48)-(3.49). The HBO trips are produced at the household, whereas the NHB trips are generated elsewhere. For the NHB trips, we can estimate how many trips are generated by the household, but we don't know (at this point in the process) where they take place.

$$\text{HBO trip productions} = P_{HBO,i} * \sum_{n=1}^N (DW_{i,n} * \text{TRIP}_{i,DW}) \quad (3.48)$$

$$\text{NHB trip productions} = P_{NHB,i} * \sum_{n=1}^N (DW_{i,n} * \text{TRIP}_{i,DW}) \quad (3.49)$$

$$\text{E-I trip productions} = \text{E-I trip attractions (see equation (3.53))} \quad (3.50)$$

where

$P_{HBO,i}$  : the percent of average daily person trips produced for HBO trips in zone i

$P_{NHB,i}$  : the percent of average daily person trips produced for NHB trips in zone i

- $DW_{i,n}$  : the number of dwelling units of household type  $n$  (e.g., single-family) in zone  $i$ .
- $TRIP_{i,DW}$  : the trip production rate of each dwelling unit of household type  $n$  in zone  $i$ , which can be derived based on TETUP's study (1994) or from survey data.

## 2) Trip attraction

Trip attractions for each zone can be developed using the following equations (3.51)-(3.53).

$$\text{HBO trip attractions} = \sigma_1 * R_i + \sigma_2 * NR_i + \sigma_3 * DW_i \quad (3.51)$$

$$\text{NHB trip attractions} = \sigma_4 * R_i + \sigma_5 * NR_i + \sigma_6 * DW_i \quad (3.52)$$

$$\text{E-I trip attractions} = C_1 + \sigma_7 * R_i + \sigma_8 * NR_i + \sigma_9 * DW_i \quad (3.53)$$

where

$R_i$  : the number of retail employees in zone  $i$ .

$NR_i$  : the number of nonretail employees in zone  $i$ .

$DW_i$  : the number of dwelling units in zone  $i$ .

$\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7, \sigma_8$  and  $\sigma_9$  are empirical calibrated coefficients.

$C_1$  : a constant.

The estimated total trips produced at the household level should be equal to the total trips attracted at the activity locations, such as workplace and shopping center. In reality, the estimation of trip productions and attractions will not be exactly equal. To bring the regional totals to balance, either the trip productions or attractions are scaled to equal regional totals (TETUP, 1994). In most cases, the regional total for trip productions, by purpose, is used as the control trip total. Because decennial census data are used to identify the household characteristics by zone, we have a greater degree of confidence in household data than we do in employment data. Thus zonal trip attractions will be balanced to match regional productions. Regional productions and attractions consist of:

- 1) Productions and attractions from internal zones.
- 2) Productions and attractions from external stations.

Because trip productions and attractions at external stations are a function of observed or forecast traffic counts at the stations, the assumption that they are fixed and not factored is acceptable. Thus trip productions and attractions at external stations will not be balanced and the difference must be balanced across the zonal trips over the study area (TETUP, 1994). The control total of trip productions for each purpose is derived using equation (3.54).

$$CT_{pi} = \sum_k P_{k,i} + \sum_e P_{e,i} - \sum_e A_{e,i} \quad (3.54)$$

where

$CT_{pi}$  : the control total of trip productions for trip purpose i.

$P_{k,i}$  : trip productions in internal zone k for trip purpose i.

$P_{e,i}$  : trip productions at external station e for trip purpose i.

$A_{e,i}$  : trip attractions at external station e for trip purpose i.

The balancing factor for each trip purpose can be computed using the following equations.

$$F_{HBO} = CT_{HBO} / \sum_i (\sigma_1^* R_i + \sigma_2^* NR_i + \sigma_3^* DW_i) \quad (3.55)$$

$$F_{NHB} = CT_{NHB} / \sum_i (\sigma_4^* R_i + \sigma_5^* NR_i + \sigma_6^* DW_i) \quad (3.56)$$

where

$F_{HBO}$  and  $F_{NHB}$  are areawide control factors.

Next step each zone's trip attractions are multiplied by the balancing factor to obtain the final balanced trip attractions. The equation is written as follows.

$$A'_{k,i} = F_i^* A_{k,i} \quad (3.57)$$

where

$A'_{k,i}$  : trip attractions in zone k for trip purpose i after balance.

$A_{k,i}$  : trip attractions in zone k for trip purpose i before balance.

External station attractions are not factored. For the home-based school (HBS) trips, they are estimated based on survey data (e.g., home interview survey).

### 3.6.3 Trip Distribution

The distributions of the home-based other (HBO), nonhome-based (NHB) and external-internal or internal-external trips are accomplished by using a gravity approach to produce trip matrices (Dickey, 1983). Furthermore, a growth-factor procedure (e.g., Fratar method) is used to estimate external-external trips because, in general, no land use data are available for areas outside the study area (Dickey, 1983).

### 3.6.4 Modal Split

Mode choice procedures attempt to estimate the total amount (or proportion) of trips on different transport modes. Since 1970, many mathematical models have been developed for this purpose (McFadden, 1978; Dickey, 1983; Ben-Akiva and Lerman, 1985). They are generally divided into two groups: aggregate and disaggregate. The former includes the trip-end and trip-interchange modal-split models. The method widely used in these models are multiple regression, cross-classification and diversion curves. The disaggregate procedures are probabilistic and behavioral in nature and are based on the theory of utility perceived by a particular traveler. Two commonly used approaches, diversion curves and the individual mode choice model, are used in the integrated land-use/transportation model. They are performed on the basis of the data available in the region being analyzed. These two procedures are described in the following two sections.

#### 1) Diversion curves

In general, this approach performs the allocation of the total trips between two competing modes of transportation according to previously determined curvilinear



diversion relationships between the modes (Dickey, 1983). The two modes are typically private automobile and public transit. The factors available for affecting mode choice processes are the characteristics of trip origin and destination (e.g., mode availability, trip purpose), the socioeconomic characteristics of tripmakers (e.g., income, auto ownership), and the operational characteristics of the competing modes (e.g., travel time and travel cost). The mode split of trips for each purpose is achieved through diversion curves that specify the percentage of trips traveling by transit, based on the ratio or difference of the impedance of the two competing modes.

The next step is to use a set of auto occupancy rates to convert auto person trips into vehicle trips. A survey data of these rates are listed in Table 3.2. The transit person trips can be converted in a similar way into vehicle trips by using a set of transit occupancy rates. Then, the total vehicle trips between each zone pair will be loaded on the network in the traffic assignment stage.

**Table 3.2 Tippecanoe County, Indiana: 1994 data on auto occupancy**

Trip purpose	Average auto occupancy
Home-based	
work	1.14
shop	1.65
other	1.65
Nonhome-based	1.44

Source: Tippecanoe County Area Plan Commission, Indiana, 1994.

There are two major weaknesses of the diversion curve approach (Dickey, 1983).

- (1) Only two modes at a time can be considered, e.g., auto and transit. It is impossible to know the percentage of trips traveling by bus, rail and streetcar. Furthermore, if a new mode were introduced, it would be difficult to determine the share of travel market it would capture.
  - (2) It is not easily transferred from one area to another.
- 2) Individual mode choice model

The mathematical formulation used in the individual mode choice model is the logit function.

$$P(i) = \exp(\mu U_i) / \sum_{j=1}^N \exp(\mu U_j) \quad (3.58)$$

where

$P(i)$  : the probability of an individual choosing mode  $i$ .

$U_i$  : the utility of mode  $i$ .

$N$  : set of mode alternatives.

$\mu$  : a scale parameter

In general, the  $U_i$  are often represented as linear functions of the variables.

$$U_i = \sum_k \beta_k X_{ik} \quad (3.59)$$

where  $\beta_k$  and  $X_{ik}$  are the coefficients and independent variables, respectively, describing the utility for mode  $i$ . These independent variables are transportation system characteristics (e.g., travel time and travel cost) and the socioeconomic characteristics of tripmakers (e.g., income, auto ownership), possibly stratified by trip purpose. (For a more detailed review of the multinomial logit and nested logit models, see Ben-Akiva and Lerman, 1985.)

The major strength of individual mode choice models is that they are sensitive to transportation policy strategies, such as transit fares, gasoline prices, and parking fees. Furthermore, they are more easily transferable from one area to another than the aggregate approaches.

It is worth noting that the logit models have an inherent disadvantage, namely the property of the independence of irrelevant alternatives (IIA). The IIA property holds that, for a specific individual, the ratio of the choice probabilities for any two alternatives is entirely unaffected by the systematic utilities of any other alternatives (Ben-Akiva and Lerman, 1985). Some situations derived from the IIA property may be unrealistic. For example, The IIA property will consider a bus system with red buses as a completely different mode from a bus system with blue buses. They compete more heavily against

each other than against the auto market. A model having the IIA property would not be able to account entirely for these differences in competitiveness.

### 3.6.5 Traffic Assignment

Traditionally, the traffic assignment procedure has been regarded as the final stage of the conventional four-stage approach (trip generation, trip distribution, modal split, traffic assignment) used to model travel demand. Origin-Destination (O-D) matrices of trips by the various modes, for different times of day or trip purposes, are required to be allocated to the specified transportation system. The traffic assignment procedures have often been used to provide the estimates of traffic flows on different highway systems, providing a basis for the comparisons and evaluations of various transportation systems proposed for short-term or long-term analysis in the future.

The two major aims of traffic assignment procedures are:

- (1) To estimate the volume of traffic on the links of the network and the turning movement at intersections.
- (2) To estimate travel costs between trip origins and destinations for use in trip distribution and mode choice stages, and in the land use allocation module.

Many factors have been identified as influencing route choice, including travel length, travel time, travel cost, knowledge of the route, queueing, age, and sex (Benshoof, 1970; Ratcliffe, 1972; Armstrong, 1977; Hutchinson et al., 1977; Lunn, 1978). However, there is no clear view as to the most appropriate criterion to use in determining route choice. Particularly in urban areas, these route choice criteria vary with the driver, the trip purpose, and the type of link considered. Minimum travel time is the most widely used criterion. This is appropriate, because it reflects the effects of congestion and is sensitive to a range of policy measures. However, a combination of travel time, travel distance and travel cost is a more appropriate criterion. In general, for any O-D pair, not all drivers will choose the same route. There may be various reasons for the differences in choices in

individuals' cases. For example, drivers traveling between any O-D pair may perceive different degrees of congestion on the same route.

Van Vliet and Dow (1976) proposed that the choice of an assignment procedure should be based on the level of congestion presented in the network. When the network is relatively uncongested, as the case in off-peak periods, stochastic choice processes may be preferable. However, with higher degrees of congestion, capacity restraint methods should be used.

Capacity restraint procedures recognize the dependence of link speeds and intersection delays on traffic flows. Without capacity restraint, many links are likely to be loaded well beyond their capacities, whereas parallel links are likely to remain lightly loaded. Under this situation, the traffic assignment is unstable, because small increases in the costs on some highly loaded links or decreases on some lightly loaded links may lead to quite different route choices. Stochastic route choice arises because drivers' perceptions of travel costs differ. From a behavioral point of view, the assumption of stochastic route choice is sound, because many empirical studies have concluded that there is no single criterion used by all drivers as the basis of route choice (Benshoof, 1970; Ratcliffe, 1972; Armstrong, 1977; Hutchinson et al., 1977; Lunn, 1978). By providing the use of several routes between each O-D pair, a stochastic route choice procedure produces a more stable assignment of traffic that is less sensitive to minor differences in link costs. The combination of stochastic route choice and capacity restraint accounts for the patterns of route choice observed in congested conditions. Variations in travel costs as perceived by individual drivers are clearly an important element in route choice procedures.

Given an O-D trip matrix representing the demand for travel, the traffic assignment problem is to determine a pattern of movements satisfying the demand that is consistent with the network system, including highways and transit, and the behavior of drivers as expressed in terms of the cost functions assumed. A solution to this problem may be regarded as a point of equilibrium between the demand for movement and the supply offered by the network system.

Two basic forms of travel behavior have been proposed as the basis for a state of network equilibrium.

- 1) Drivers choose their routes independently, in their own best interest, on the basis of the traffic conditions resulting from the choices of others -- User Optimization.
- 2) Drivers cooperate in their choice of routes so as to produce a pattern of traffic flows giving maximum benefit to the whole network system -- System Optimization.

In 1952, Wardrop proposed two principles of route choice in accordance with these two forms of behavior.

- 1) The costs on all routes used between any given pair of O-D are equal and are not greater than the cost experienced by a single vehicle on any unused route between them.
- 2) The average journey cost over all routes used is as small as possible.

In practice, Wardrop's first principle is the most likely basis for network equilibrium. Most of the theoretical studies have attempted to produce a solution in accordance with this principle. Beckmann et al. (1956) showed that if the cost,  $C_a$ , on any link  $a$ , is a function of the flow,  $V_a$ , on the link. The cost function is an increasing function of  $V_a$ , and the flows satisfying Wardrop's user-optimized equilibrium principle are unique. Dafermos (1972) showed how different types of vehicles, drivers with different route choice criteria, and users in different parts of the network may be simultaneously assigned. Fisk (1980) proposed an equilibrium assignment based on the logit model of route choice (Dial, 1971) that allows drivers to have different perceptions of cost on the network links. Smith (1979) proposed a less restrictive principle of route choice, in which drivers only change routes when their total travel costs, based on today's route costs, are reduced. It implies that all possible route choices of tomorrow's traffic flows may be determined on the basis of today's route travel costs. Equilibrium assignment approaches produce link flow patterns in accordance with Wardrop's first principle. It has been shown that the Frank-Wolfe algorithm converges in a stable manner, i.e., estimates of flows from successive

iterations do not oscillate but continuously converge, towards the equilibrium solution (Evans, 1976; Van Vliet and Dow, 1979).

In order to estimate the accessibility, or generalized cost of travel time and travel cost, between each O-D pair, three assignment methods (all-or-nothing, equilibrium, stochastic) are currently used in the integrated land-use/transportation model. They are performed on the basis of the characteristics of the region being analyzed. These assignment procedures are described below. Let the origin-destination matrix be denoted by  $\mathbf{q}$  with entries  $q_{rs}$ . In other words,  $q_{rs}$  is the trip rate between origin  $r$  and destination  $s$  during the period of analysis. And, let  $x_a$  and  $t_a$  represent the flow and travel time, respectively, on link  $a$  of a network being studied. Furthermore,  $t_a = t_a(x_a)$ , where  $t_a(\cdot)$  represents the relationship between flow and travel time for link  $a$ . In other words,  $t_a(x_a)$  is the link performance function.

#### 1) All-or-Nothing assignment

In an All-or-Nothing (A-O-N) procedure, the trips from any origin zone to any destination zone are loaded into a single minimum cost path between them. This is a free-flow assignment with no consideration of link capacities or link types.

#### 2) Equilibrium assignment

The equilibrium assignment problem is to find the link flows that satisfy the user-equilibrium (UE) criterion when all the origin-destination trips have been appropriately assigned. Two equilibrium assignment techniques are currently used, based on the characteristics of the region being analyzed.

##### (1) Capacity restraint assignment

As the A-O-N assignment, all trips between each O-D pair are loaded on the minimum cost path between them. The link time is adjusted according to the user-specified volume/capacity time adjustment curve or the standard Bureau of Public Roads capacity restraint formula.

$$T_n = T_{n-1} * \left[ 1.0 + 0.15 * \left( \frac{\text{Assigned Volume}}{\text{capacity}} \right)^4 \right] \quad (3.60)$$

where n is the current restraint iteration.

One example of the steps of the capacity restraint assignment procedure is represented as follows.

Step 0. Initialization

A-O-N assignment based on  $t_a^0 = t_a(0)$ ,  $\forall a$ .

$\Rightarrow$  obtain a set of link flows  $\{x_a^0\}$ , set  $n = 1$ .

Step 1. Update

Set  $t_a^n = t_a(x_a^{n-1})$ ,  $\forall a$ .

Step 2. Smoothing

Set  $t_a^n = 0.75t_a^{n-1} + 0.25t_a^n$ ,  $\forall a$ .

Step 3. Network loading

Assign all trips to the network using A-O-N based on travel time  $\{t_a^n\}$

$\Rightarrow$  obtain link flow  $\{x_a^n\}$ .

Step 4. Stopping rule

If  $n = N$ , go to step 5, otherwise,  $n = n+1$ , go to step1.

Step 5. Averaging

Set  $x_a^* = 1/4 \sum_{l=0}^3 x_a^{n-l}$ ,  $\forall a$ , and stop.

(2) Incremental assignment

Another method for attaining the UE solution that assigns a portion of O-D trip matrix at each iteration is called incremental assignment. For each iteration, a user-specified percentage of trips between O-D pairs is loaded on the minimum cost path determined during path building. The link time is adjusted according to the user-specified



volume/capacity speed adjustment curve or the above standard Bureau of Public Roads capacity restraint formula. Its steps are written below.

Step 0. Preliminaries

Divide trips between each O-D pair into N equal portions (i.e., set  $q_{rs}^n = q_{rs}/N$ ).

Set  $n=1$  and  $x_a^0 = 0, \forall a$ .

Step 1. Set  $t_a^n = t_a(x_a^{n-1}), \forall a$ .

Step 2. Incremental loading

Perform A-O-N assignment based on  $\{t_a^n\}$ , but using only the trip rate  $q_{rs}^n$  for each O-D pair  $\Rightarrow$  yield  $\{W_a^{(n)}\}$ .

where  $W_a^{(n)}$  is the trips loaded on link a for iteration n.

Step 3. Flow summation.

$$x_a^n = x_a^{n-1} + W_a^{(n)}, \forall a.$$

Step 4. Stopping rule

If  $n = N$ , stop, otherwise, set  $n = n+1$  and go to step 1.

3) Stochastic assignment

The stochastic assignment performs a probabilistic multipath traffic assignment. Trips are assigned to all acceptable paths between O-D pairs. The multinomial logit model is widely used in the stochastic multiple route choice procedures. The probability of the route R chosen by drivers can be estimated by using the following formula.

$$\Pr(R) = \exp(\theta V_{R,t}) / \sum_{s \in C} \exp(\theta V_{s,t}) = \exp(-\theta Cr) / \sum_{s \in C} \exp(-\theta Cs) \quad (3.61)$$

where Cs: the perceived cost on route s.

Cr: the perceived cost on route r.

$\theta$  : a scale parameter

The stochastic route choice procedure adopted in this research is the Dial model. The Dial model evaluates route by comparing the total costs between trip origin and destination on each of the acceptable routes. The probability given in (3.61) is the basis



for route choice in the model proposed by Dial (1971). Basically, trips are allocated to any efficient route R so that,

$$V_{R,t} = V^* \exp(-\theta (C_R - C^*)) = V^* \exp(-\theta \Delta R) \quad (3.62)$$

where  $\Delta R = C_R - C^*$  is the “excess” cost in using route R rather than the minimum cost route, the cost on which is  $C^*$ ,  $V_{R,t}$  and  $V^*$  are the flows on route R and the minimum cost route respectively.  $\theta$  is a user-specified dispersion parameter.

As mentioned in Section 3.6.4, an inherent disadvantage of the logit model is the IIA property. Sheffi (1985) mentioned that the deficiency of the logit model is not unique to mode-choice problems. This limitation of the IIA property may exhibit some unreasonable results in a logit route-choice problem. The route-choice anomaly cited by Sheffi stems from the logit’s inability to account for the correlation between the perceived travel times of the various routes. (Sheffi, 1985, pp.294-296)

Another possible anomaly that may arise in a logit route-choice procedure is that the path choice probabilities are determined on the basis of travel time differences (see equation (3.62)). In fact, the distribution of the perceived path travel time should be a function of the lengths of the alternative routes. The longer routes should have larger variance of perception. However, a logit route-choice model can not identify such a characteristic (Sheffi, 1985).

### 3.7 Model Calibration

Calibration of the residential location model and the employment location model involves: (a) specification of appropriate indicators of *Goodness-of-Fit* of the models to the data, (b) specification of mathematical equations (e.g., equations (3.4)-(3.7)) to capture the real relationships among economic activities, and (c) adjustments to parameters in the equations so that the estimated values of variables (e.g., population and mean trip length) equal the observed ones. The equations in the residential and the

employment location models are intrinsically nonlinear, and the data from which their parameters must be estimated are not normally distributed (Putman, 1996). The appropriate goodness-of-fit measure for calibration of the residential and the employment location models is the likelihood function, analogous to the notion of maximum likelihood as developed in econometrics (Pindyck and Rubinfeld, 1991). The likelihood function for goodness-of-fit has the following form:

$$L = \sum_i V_i \ln \hat{V}_i \quad (3.63)$$

where  $L$  is the total likelihood measure,  $V_i$  is the observed value of the dependent variable, and  $\hat{V}_i$  is the estimated value (Putman, 1983). In the residential location model, the dependent variable would be population or households of a particular type (e.g., single-family, two-family or multi-family) located in a particular zone, and in the employment location model, the dependent variable would be the number of employees of a particular type (e.g., retail or nonretail) located in a particular zone.

The meaning of “best fit” is that the difference between the estimated values of the dependent variable and the observed values in the calibration data set is minimized (Putman, 1983 and 1996). A perfect fit means that the estimated  $\hat{V}_i$  equals the observed  $V_i$  for each zone  $i$ . The value of the best fit of the likelihood function would be:

$$L_b = \sum_i V_i \ln V_i \quad (3.64)$$

Because of likelihood functions (3.63) and (3.64) and because the control total of the dependent variable within each time period must be a known constant (i.e.,  $\sum_i V_i = \sum_i \hat{V}_i$ ), a likelihood function of goodness-of-fit can be redefined. The likelihood function takes the following form (Putman, 1983).

$$\psi = L - L_b = \sum_i V_i \ln \hat{V}_i - \sum_i V_i \ln V_i \quad (3.65)$$

Because  $V_i \geq 0$ , the range of the likelihood  $\psi$  is from  $-\infty$  to 0 (for a perfect fit). The calibration process for the residential and the employment location models uses an unconstrained nonlinear maximization algorithm to maximize the value of  $\psi$  with respect to the parameters  $\beta_1^n$ ,  $\beta_2^n$ ,  $\alpha_1$ ,  $\alpha_2, \dots$ , and  $\alpha_n$  (see equation (3.4)). The above calibration procedures are the same as those described in the DRAM and EMPAL models developed by Putman (1983).

A second criterion for goodness-of-fit in the calibration of the residential location model must ensure that the model's mean work-to-home trip cost is equal to the observed mean work-to-home trip cost (Baxter and Williams, 1975). This is only possible when information about the cost of travel is available. Based on the presence of mean trip cost and the criterion of likelihood function  $\psi$  (i.e., equation (3.65)), a possible solution for calibrating the model is to accept that the estimated mean work-to-home trip cost ( $\bar{C}_{\text{mod}P}$ ) cannot be measured beyond a certain degree of accuracy, and then set up a range within which the estimated mean trip cost must lie. The optimal solution for the parameters  $\beta_1^n$ ,  $\beta_2^n$ ,  $\alpha_1$ ,  $\alpha_2, \dots$ , and  $\alpha_n$  (see equation (3.4)) within this range can then be found to satisfy the second goodness-of-fit criterion (i.e., constraint (3.67)). The residential location model would be run to adjust values of  $\beta_1^n$ ,  $\beta_2^n$ ,  $\alpha_1$ ,  $\alpha_2, \dots$ , and  $\alpha_n$  to find

$$\max \psi = \sum_i P_i \ln \hat{P}_i - \sum_i P_i \ln P_i \quad (3.66)$$

$$\text{subject to } \bar{C}_{\min P} \leq \bar{C}_{\text{mod}P} \leq \bar{C}_{\max P} \quad (3.67)$$

where

- $P_i$  : the observed population living in the zone  $i$ .
- $\hat{P}_i$  : the estimated population living in the zone  $i$  by the integrated model.
- $\bar{C}_{\max P}$  : the upper limit of mean work-to-home trip cost.
- $\bar{C}_{\text{mod}P}$  : the estimated mean work-to-home trip cost.
- $\bar{C}_{\min P}$  : the lower limit of mean work-to-home trip cost.

Likewise, the second goodness-of-fit calibration criterion for the employment location model also needs to ensure that the estimated mean home-based shopping trip cost is equal to the observed mean home-based shopping trip cost. The optimal solution for the parameters  $\beta_3^n$ ,  $\beta_4^n$ ,  $\gamma_1$ ,  $\gamma_2, \dots$ , and  $\gamma_k$  (see equation (3.13)), within acceptable limits, can then be solved to satisfy the second goodness-of-fit criterion (i.e., constraint (3.69)). In a similar way, the employment location model would be run to adjust values of  $\beta_3^n$ ,  $\beta_4^n$ ,  $\gamma_1$ ,  $\gamma_2, \dots$ , and  $\gamma_k$  to find

$$\max \psi = \sum_j E_j \ln \hat{E}_j - \sum_j E_j \ln E_j \quad (3.68)$$

$$\text{subject to } \bar{C}_{\min} e \leq \bar{C}_{\text{mod}} e \leq \bar{C}_{\max} e \quad (3.69)$$

where

- $E_j$  : the observed employees working in zone j.
- $\hat{E}_j$  : the estimated employees working in zone j by the integrated model.
- $\bar{C}_{\max} e$  : the upper limit of mean home-to-shop trip cost.
- $\bar{C}_{\text{mod}} e$  : the estimated mean home-to-shop trip cost.
- $\bar{C}_{\min} e$  : the lower limit of mean home-to-shop trip cost.

Another goodness-of-fit calibration criterion, RMSE (root-mean-square error) (Pindyck and Rubinfeld, 1991), will be used to compare the estimated and observed population and employees in the study area. In the residential location model,

$$\text{RMSE}_{\text{pop}} = \sqrt{(1/N) \sum_{i=1}^N (P_i - \hat{P}_i)^2}$$

where

- $\text{RMSE}_{\text{pop}}$  : the root-mean-square error for population.
- $P_i$  : the observed population living in the zone i.
- $\hat{P}_i$  : the estimated population living in the zone i by the integrated model.
- $N$  : the number of zones in the study area.

In the employment location model,

$$\text{RMSE}_{emi} = \sqrt{(1/N) \sum_{j=1}^N (E_j - \hat{E}_j)^2}$$

where

$\text{RMSE}_{emi}$  : the root-mean-square error for employment.

$E_j$  : the observed employees working in the zone i.

$\hat{E}_j$  : the estimated employees working in the zone i by the integrated model.

N : the number of zones in the study area.

#### Gradient search method

The approach used to calibrate parameters in the residential and employment location models is gradient search technique (Bazaraa et. al., 1993). This procedure is appropriate to the nonlinear equations of the two models, and to the non-normal distribution of the data.

In the residential location model, equation (3.4) can be rewritten in the following general form.

$$P_i = \sum_j E_j B_j w_i^\alpha t_{ij}^{\beta_2} \exp(-\beta_1 t_{ij}) \quad (3.70)$$

and

$$B_j = [\sum_i w_i^\alpha t_{ij}^{\beta_2} \exp(-\beta_1 t_{ij})]^{-1} \quad (3.71)$$

where

$P_i$  : the population living in zone i.

$E_j$  : the number of employees working in zone j.

$t_{ij}$  : the disutility (travel time or travel cost or both) of travel between zone i and j.

$w_i$  : the residential attractiveness, such as vacant available land, the number of households or dwelling units, the average monthly rent of housing units, for population living in zone i.

$\beta_1, \beta_2, \alpha$  : empirically calibrated parameters.

The term  $B_j w_i^\alpha t_{ij}^{\beta_2} \exp(-\beta_1 t_{ij})$  represents the average probability of a commuter employed at zone j will choose to live in zone i.

The calibration processes used in the residential location model to maximize the value of  $\psi$  (i.e., equation (3.66)) with respect to the parameters  $\beta_1$ ,  $\beta_2$  and  $\alpha$  are derived using the following steps.

**Step 0: Initialization**

Choose the starting values for  $\beta_1(t)$ ,  $\beta_2(t)$  and  $\alpha(t)$ , pick stopping tolerance  $\varepsilon > 0$ , and set  $t = 0$ .

**Step 1: Gradient**

Compute the gradient,  $\nabla \psi(x(t)) = (\partial \psi / \partial \beta_1, \partial \psi / \partial \beta_2, \partial \psi / \partial \alpha)$ , of the likelihood function  $\psi(x(t)) = \psi(\beta_1(t), \beta_2(t), \alpha(t))$ , and its norm  $\|\nabla \psi(x(t))\| = \sqrt{\sum_j (\partial \psi / \partial x_j)^2}$ .

By the chain rule for the derivative, the gradient of  $\psi$  with respect to  $\beta_1$  is derived:

$$\begin{aligned} \frac{\partial \psi}{\partial \beta_1} &= \frac{\partial \psi}{\partial \hat{P}_i} \frac{\partial \hat{P}_i}{\partial \beta_1} = \sum_i \frac{P_i}{\hat{P}_i} \frac{\partial \hat{P}_i}{\partial \beta_1} \\ &= \sum_i \frac{P_i}{\hat{P}_i} \left[ \sum_j \left( E_j B_j A_{ji} \right) \left( B_j \left( \sum_i t_{ji} A_{ji} \right) - t_{ji} \right) \right] \end{aligned} \quad (3.72)$$

where

$$A_{ji} = w_i^\alpha t_{ij}^{\beta_2} \exp(-\beta_1 t_{ij}) \quad (3.73)$$

In a similar way, the  $\partial \psi / \partial \beta_2$  and  $\partial \psi / \partial \alpha$  can be derived:

$$\frac{\partial \psi}{\partial \beta_2} = \frac{\partial \psi}{\partial \hat{P}_i} \frac{\partial \hat{P}_i}{\partial \beta_2} = \sum_i \frac{P_i}{\hat{P}_i} \frac{\partial \hat{P}_i}{\partial \beta_2}$$

$$= \sum_i \frac{P_i}{\hat{P}_i} \left[ \sum_j \left( E_j B_j A_{ji} \right) \left( \ln t_{ji} - B_j \sum_i (\ln t_{ji}) A_{ji} \right) \right] \quad (3.74)$$

and

$$\begin{aligned} \frac{\partial \psi}{\partial \alpha} &= \frac{\partial \psi}{\partial \hat{P}_i} \frac{\partial \hat{P}_i}{\partial \alpha} = \sum_i \frac{P_i}{\hat{P}_i} \frac{\partial \hat{P}_i}{\partial \alpha} \\ &= \sum_i \frac{P_i}{\hat{P}_i} \left[ \sum_j \left( E_j B_j A_{ji} \right) \left( \ln w_i - B_j \sum_i (\ln w_i) A_{ji} \right) \right] \end{aligned} \quad (3.75)$$

Step 2: Stationary point

If  $\| \nabla \psi(x(t)) \| < \varepsilon$ , then stop. Otherwise, go to Step 3.

Step 3: Line search

Use golden section method (Bazaraa et. al., 1993) to do one-dimension line search, finding a  $\lambda$  to

$$\max \psi(x(t) + \lambda (\nabla \psi(x(t)) / \| \nabla \psi(x(t)) \|))$$

Step 4: Advance

$$\text{Update } x(t+1) = x(t) + \lambda (\nabla \psi(x(t)) / \| \nabla \psi(x(t)) \|),$$

then  $t = t + 1$  and return to Step 1.

The calibration process of parameters in the employment location model uses the same four-step procedure as above.

Asymptotic t-Test in model calibration

The estimated parameters in the residential and employment location models are maximum likelihood estimators. Maximum likelihood estimators are asymptotically

normal and asymptotically efficient, and their asymptotic variance is given by the Cramer-Rao bound (Ben-Akiva and Lerman, 1985).

$$\text{Var}(\beta) \geq (-E[\partial^2 \psi / \partial \beta^2])^{-1} \quad (3.76)$$

Equation (3.76) is called the Cramer-Rao bound, where  $\beta$  is an empirically estimated parameter,  $\text{Var}(\beta)$  is the variance of  $\beta$ , and  $\partial^2 \psi / \partial \beta^2$  is the second partial derivative of  $\psi$ .

A covariance matrix with respect to the estimated parameters in the residential and employment location models can be calculated using second order partial derivatives. These derivatives are derived in the parameter estimation procedure of the gradient search method and can be used to compute the asymptotic t-statistics. The asymptotic t-Test is used to indicate the significance of each parameter in the equations of the residential and employment location models. This will enable us to identify which parameter is appropriate to incorporate in the residential and employment location models. An example of this asymptotic t-Test will be conducted in Section 5.1.1.



**Table 3.1 Lot area (square feet) per dwelling unit, by district**

Kind of Dwelling	R1	R1A	R1B	R2	R3	R4
Single-family with community sewage disposal system	10000	7500	6000	7500	6000	6000
Single-family with individual sewage disposal system ***	15000	15000	15000	15000	15000	15000
Two-family with community sewage disposal system	----	----	----	3750	3000	3000
Two-family with individual sewage disposal system ***	----	----	----	10000	10000	10000
Multi-family with community sewage disposal system	----	----	----	----	* 2000	* 2000
Multi-family with individual sewage disposal system ***	----	----	----	----	** 3500	** 3500

R1: Residential zones for low density single-family housing

R1A: Residential zones for low and medium density single-family housing

R1B: Residential zones for low, medium and relatively high density single-family housing

R2: Residential zones for medium density single-family and two-family housing

R3: Residential zones for relatively high density single-family, two-family and multi-family housing and mobile home parks

R4: Residential zones for single-family, two-family housing, and higher density multi-family housing

\* : This number applies for each of the first three dwelling units. Add 1000 for each additional unit.

\*\* : This number applies for each of the first three dwelling units. Add 1500 for each additional unit.

\*\*\* : Or as prescribed by the Indiana State Board of Health requirement for individual sewage disposal systems.

**Table 3.1 (continued)**

Kind of Dwelling	AB	LB	GB	CB	I	A	AA
Single-family with community sewage disposal system	6000	6000	6000	----	10000	15000	15000
Single-family with individual sewage disposal system ***	15000	15000	15000	----	15000	15000	15000
Two-family with community sewage disposal system	3000	3000	3000	3000	5000	7500	----
Two-family with individual sewage disposal system ***	10000	10000	10000	10000	10000	15000	----
Multi-family with community sewage disposal system	* 2000	* 2000	* 1500	* 800	----	----	----
Multi-family with individual sewage disposal system ***	** 3500	** 3500	** 3500	** 3500	----	----	----

AB: Accommodation business zone

LB: Business zones for local commercial activity

GB: General business zones for all kinds of commercial activity, including retailing, services, some wholesaling, self-storage warehousing and contracting

CB: Central business zones in the central parts of the cities for retailing, service, government and professional offices, cultural, recreational and entertainment establishments, housing, commercial lodging and transportation facilities

I: Industrial zones for industrial activity

A: Agricultural zones for agricultural use and only limited residential development where little or no urbanization has occurred or is planned to occur

AA: Select agricultural zones having either highest or a combination of high and highest potential for agricultural use, where no urbanization has occurred, and which are to be reserved for the wide range of agricultural activities

\* : This number applies for each of the first three dwelling units. Add 1000 for each additional unit.

\*\* : This number applies for each of the first three dwelling units. Add 1500 for each additional unit.

\*\*\* : Or as prescribed by the Indiana State Board of Health requirement for individual sewage disposal systems.

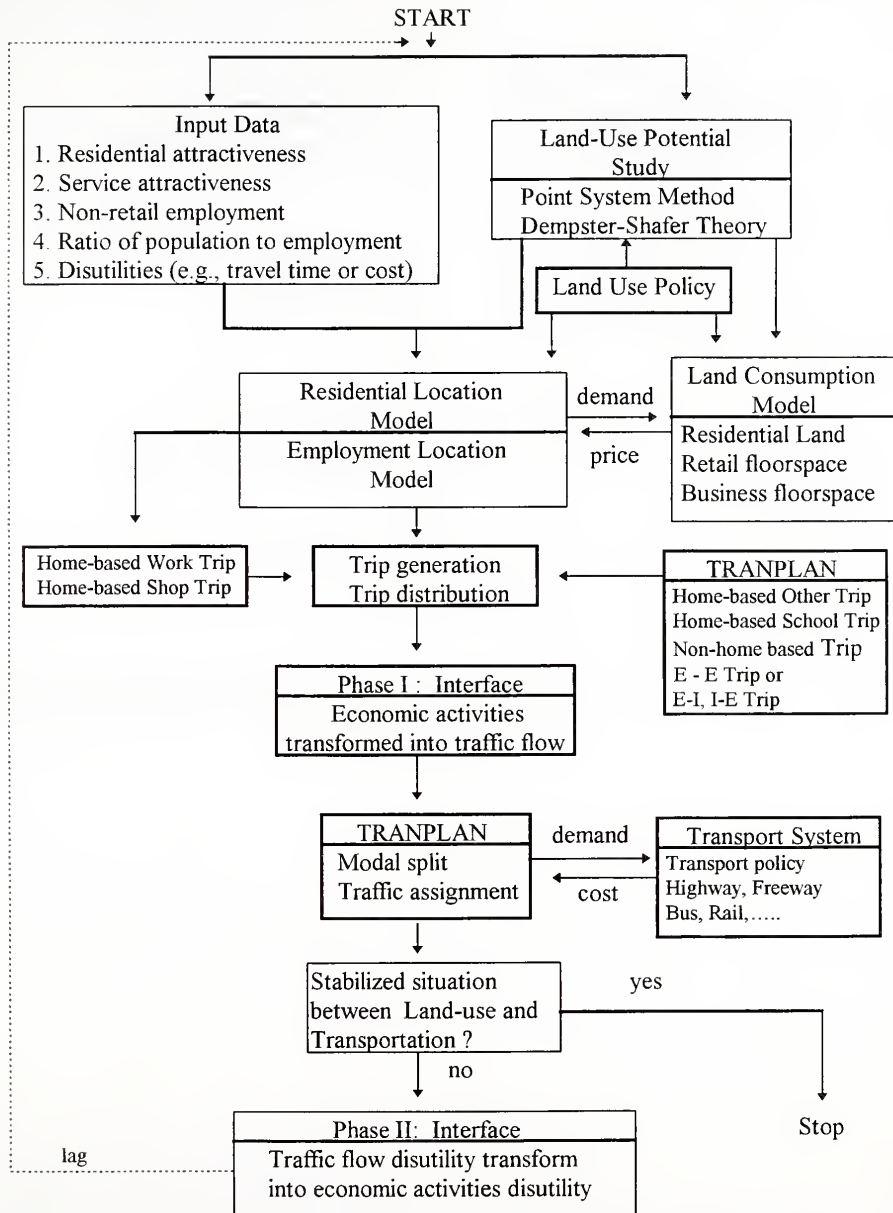
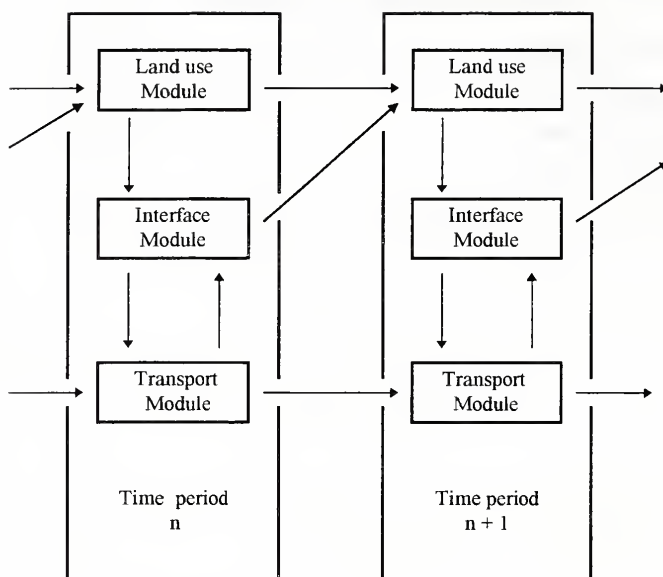
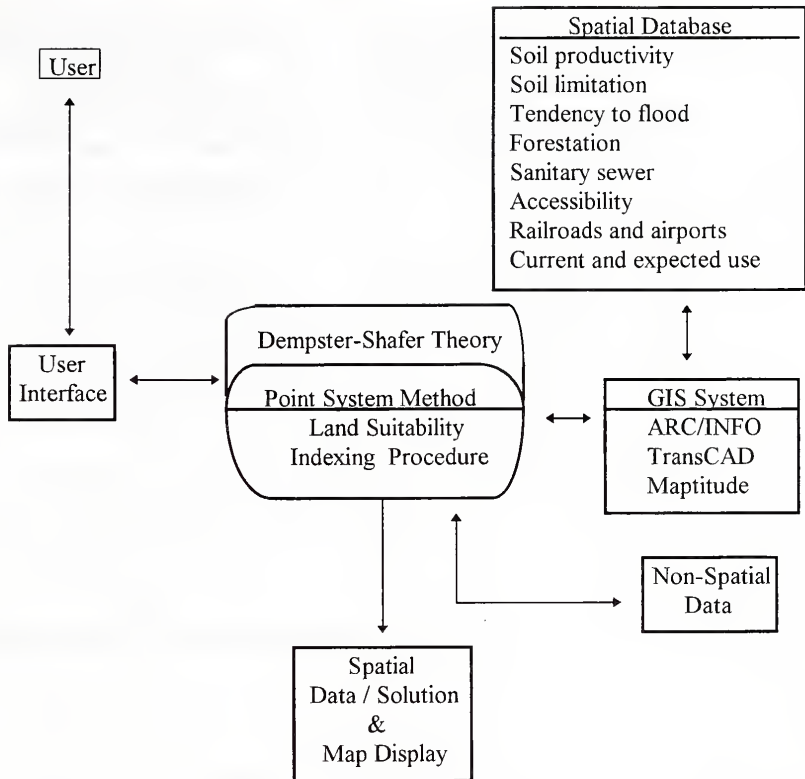


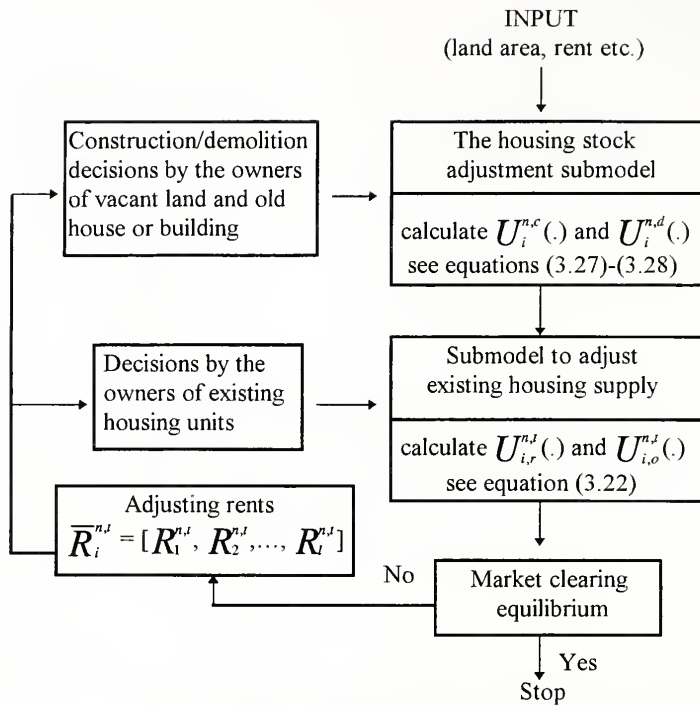
Figure 3.1 Framework of Land-Use/Transportation Model



**Figure 3.2 Order of model operation for simulating the dynamics of interaction between Land use and Transportation [Echenique et al., 1990]**



**Figure 3.3 Structure of land use potential model**



**Figure 3.4 The framework of land consumption model**

## CHAPTER 4 DATA REQUIREMENT AND ANALYSIS

### 4.1 Data Requirement

The data required for land use allocation module consist of the following items:

- 1) Households (or population) and dwelling units are collected by categories, including single-family, two-family, and multi-family (or high, middle, and low income groups) by zonal level. Employment is divided into several categories (for example, retail, manufacturing and other employment, or basic and nonbasic employment) by zonal level. Total land area for each zone in the study area is calculated from the land use potential study.
- 2) Ratio of population in household type  $n$  (e.g., single-family, two-family and multi-family) who are employees in employment type  $k$  (e.g., basic and nonbasic employment) to the total number of employees in employment type  $k$  over the study area. (For example, if there are 100 persons living in the single-family house who are also employees of basic employment in the study area and, at the same time, the total number of employees of basic employment over the study area is 200, then  $\alpha_{k,n} = 100/200 = 0.5$ .)
- 3) Ratio of the number of employees in employment type  $k$  to the sum of population and employees over the study area. (For example, if there are 800 persons living in the study area and 200 employees working there at the same time, and if there are 240 employees in basic employment in this area, then  $\delta_k = 240/(800+200) = 0.24$ .)

The ratios in items 2 and 3 above are calculated by using data collected in item 1.

- 4) Amount of land that can be used for residential development by zonal level.
- 5) Amount of land that can be used for retail (or nonbasic) and nonretail (or basic) development by zonal level.

The data for items 4 and 5 are calculated from the results of the land use potential study.

- 6) Average area (or floorspace) occupied by each retail (or nonbasic) employee in the study area or by zonal level.
- 7) Average area (or floorspace) occupied by each nonretail (or basic) employee in the study area or by zonal level.

The data for items 6 and 7 are collected by survey or previous study.

- 8) The locations and number of nonretail (or basic) employment for the forecast year in the study area are determined outside the integrated model. In fact, data on nonretail (or basic) employment can be easily collected by each MPO.
- 9) Generalized cost (e.g., a combination of travel time and travel cost) between each O-D zone pair is estimated after traffic assignment procedures in the travel demand module. The generalized cost will be fed back to the land use allocation module for estimating the spatial patterns of population and employment in the study area.

In the land use potential study, information about the eight factors or evidences (soil productivity, soil limitation, tendency to flood, forestation, sanitary sewer availability, road accessibility, railroads and airports, and current and expected use) needs to be collected before calculating the land area of residential use and employment (nonretail and retail). The database for the eight factors was built based on a raster-based structure in TransCAD, which is a commercial GIS package. The process of data assembling in the land use potential study for Tippecanoe County, Indiana is described in the following paragraphs.

Data for the three soil-related factors (soil productivity, soil limitation, and tendency to flood) are based on soil survey maps (Soil Survey, 1959). Information on soil productivity is gathered from Soil Survey. A measure of general soil productivity, ranging from low to very high, is given for each of 179 soil types (Soil Survey, pp. 100-117). Data about relative levels of limitation (slight, moderate and severe) for “dwellings without basements” and “small commercial buildings” have been collected from Soil Survey and USDA-SCS-Indiana Soil Survey Interpretation Sheets (1975). The Soil Survey Interpretation Sheets also provides information on each soil’s tendency to flood. Alluvial and organic soils are generally found in the flood plains adjacent to rivers and streams.



Other flooding soils are upland soils having a tendency to experience ponding. These flooding soils are not suitable for residential, commercial, and industrial uses or highway construction.

Forestation indicates the presence or absence of significant numbers of trees. A GIS database for forestation can be transformed from a series of aerial photographs. In Figure 4.1, “Yes” means that at least 50 percent of the area in a cell has significant numbers of trees; “No” represents that less than 50 percent of the area in a cell has significant numbers of trees.

Figure 4.2 shows trunk lines with excess capacity in that portion of the Tippecanoe County where sites already are being served by the trunk lines, and indicates distances -- within 1000' (i.e., 0'-1000'), within 0.5 mile (i.e., 1001'-2640'), within 1.5 miles (i.e., 2641'- 7920', and outside 1.5 miles (i.e., > 7921') -- from trunk lines for land adjacent to, but not yet served by those trunk lines.

Accessibility is a measure of proximity to major and minor roadways and their intersections. Four levels of proximity are included to describe the accessibility of each zone. They are:

1. Distances from a paved road are less than 0.5 mile.
2. Distances from a paved road are greater than 0.5 mile.
3. At a major intersection or along state/federal right-of-way, and
4. Along interstate right-of-way.

“Major intersection” accessibility means that the centroid of a cell is within a 0.25-mile radius of either the intersection of two state or federal highways or one state road and one federal highway, or the intersection of an interstate highway entry or exit ramp with another roadway. A graphic GIS representation of the Tippecanoe County road system, which is extracted from Census TIGER/Line Files (U.S. Department of Transportation), is shown in Figure 4.3.

The factor “railroads and airports” provides a distance measure to these major transportation facilities and their areas of influence. Figure 4.4, which is also extracted

from Census TIGER/Line Files, shows a display of the location of railroads and airports in Tippecanoe County, Indiana.

The factor “current and expected use” represents a generalized indication of how land is being used at the present time. A graphic representation of “current and expected land use” is shown in Figure 4.5. The symbols A, C, I, O, and R represent agricultural, commercial, industrial, open space and residential use, respectively.

In the land consumption model, the following data need to be collected.

- 1) Average monthly rents of housing units by zonal or township level, which are obtained from census data.
- 2) Land area available for residential use in each zone, which is calculated based on the results of the land use potential study.
- 3) The amount of housing units available for renter occupancy, the amount of housing units available for owner occupancy, and the number of housing units remaining vacant in each zone or township at each time period. These data can be obtained from the census.
- 4) The amount of new housing constructed in zone  $i$  during each time period and the amount of old housing demolished in zone  $i$  during each time period. In general, these data can be collected by each MPO.
- 5) The number of dwellings of household type  $n$  (e.g., single-family, two-family and multi-family housing) in zone  $i$  aging into the over thirty years category during each time period. This type of data can be calculated from a simple cohort-survival model based on census data.

In the travel demand module, the data required for trip generation and trip distribution stages consist of:

- 1) Total population in the study area and population in each zone within the study area.
- 2) The number of dwelling units or housing units in each zone.
- 3) Household size -- (population in zone  $i$ )/(the number of dwelling units in zone  $i$ ) -- by zonal level.

- 4) Auto ownership -- (the number of cars in zone i)/(the number of dwelling units in zone i) -- by zone.
- 5) The number of retail and nonretail employees in each zone.
- 6) Survey data for trip production rates (e.g.,  $TRIP_{i,DW}$  in Section 3.6.2) by zone.
- 7) Survey data for trip attraction rates by zone. (i.e., equations (3.51)-(3.53) in Section 3.6.2.)
- 8) Friction factor survey data over the study area.

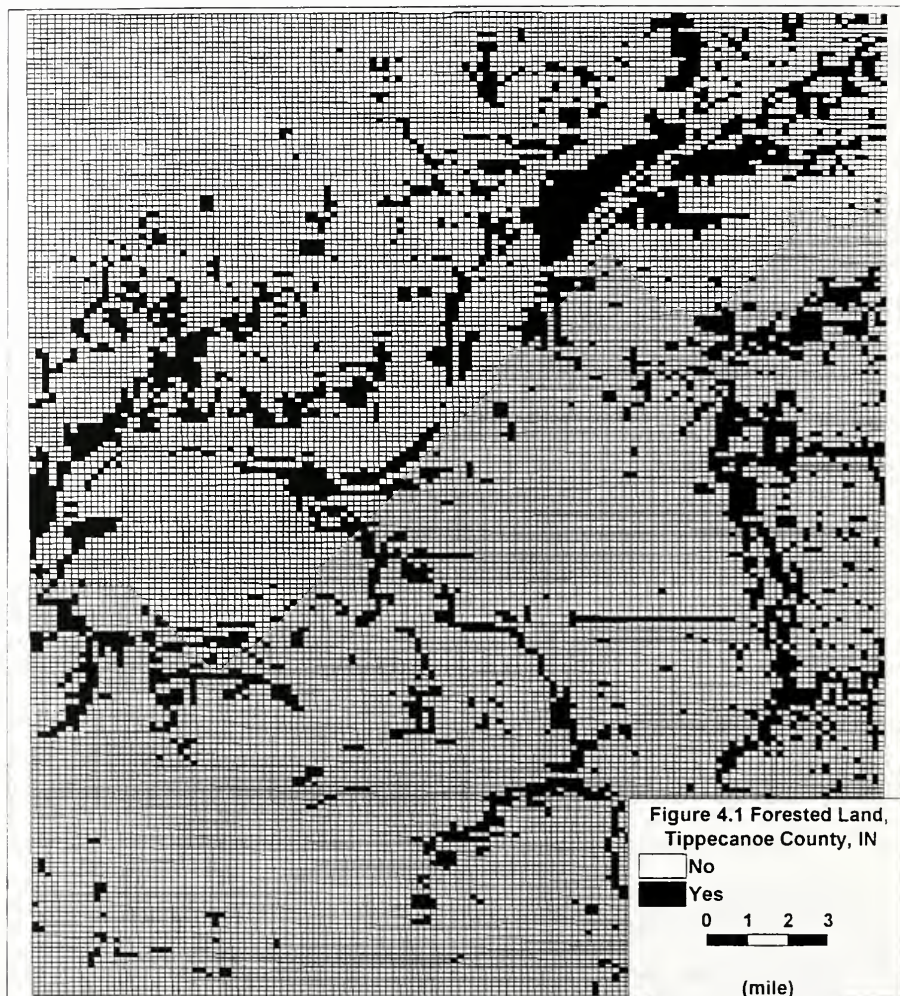
All above data can be collected from the census or can be assembled by each MPO.

In the modal split stage, a set of auto occupancy rates that converts auto person trips into vehicle trips for each trip purpose is necessary. An example of survey data for these rates is shown in Table 3.2. Likewise, a set of transit occupancy rates that transforms transit person trips into vehicle trips is also necessary. Then, the total vehicle trips between each O-D pair will be loaded on the network in the traffic assignment stage.

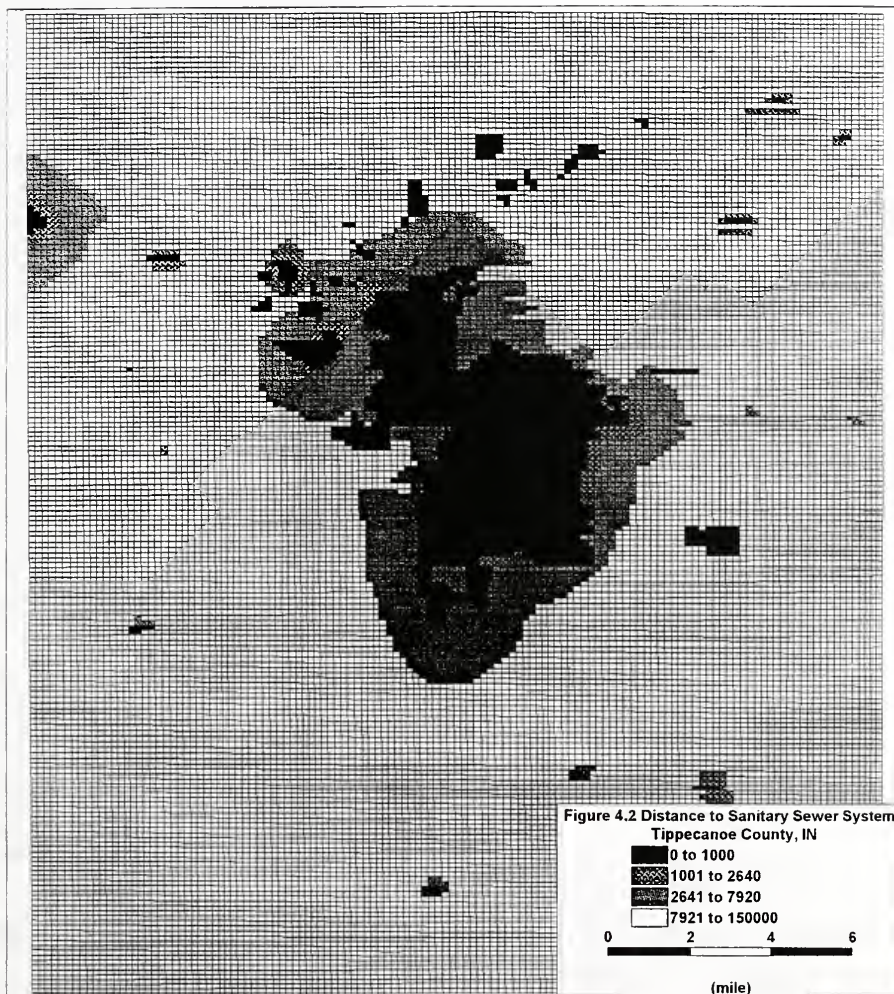
The transportation network used in the traffic assignment procedure is an abstraction of the actual transport system that may contain one or several modes (e.g., automobile, bus, and rail). The process of preparing a 'coded' network is time-consuming. The coded network will contain links, which represent all the connections between O-D pairs, as well as nodes that show the connection of tripmakers (e.g., residents and employees) in each zone to the network. Data usually required for each link are:

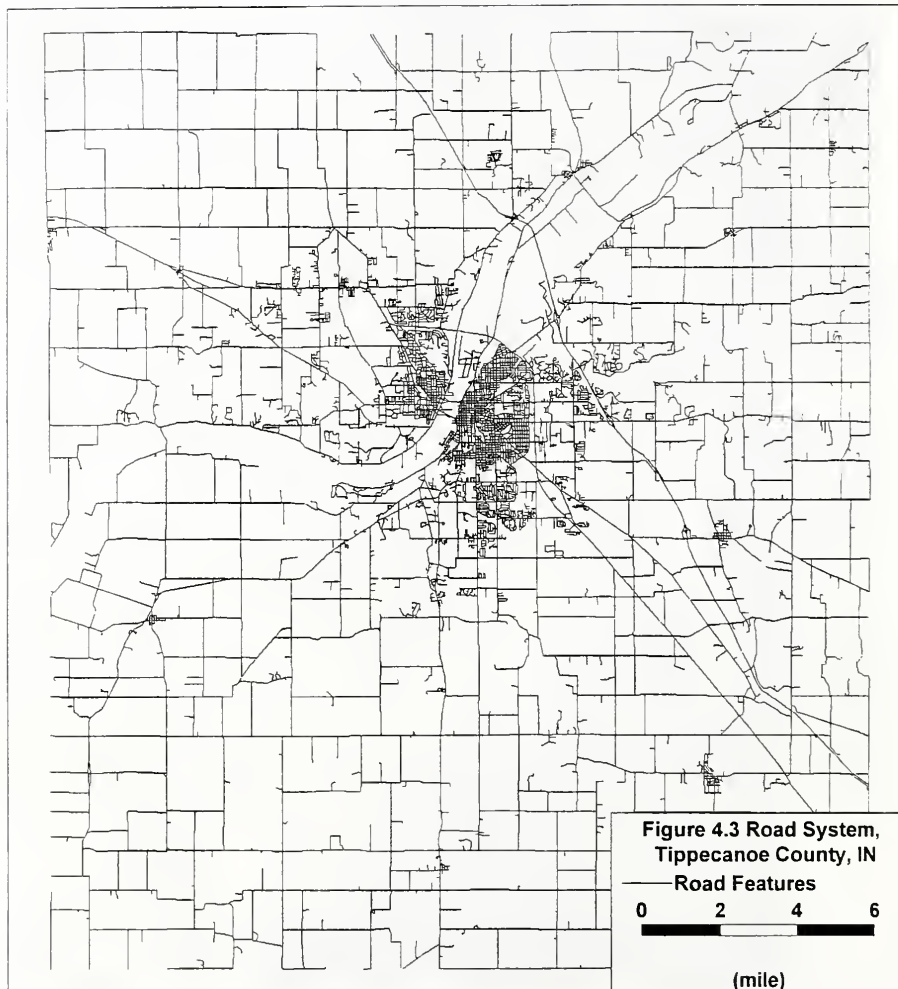
- 1) Link length.
- 2) Link type: this conveys information about width or number of lanes, speed/flow relationship, and capacity.
- 3) Exit type: this indicates the nature of the intersection (e.g., roundabout) and the intersection capacity.

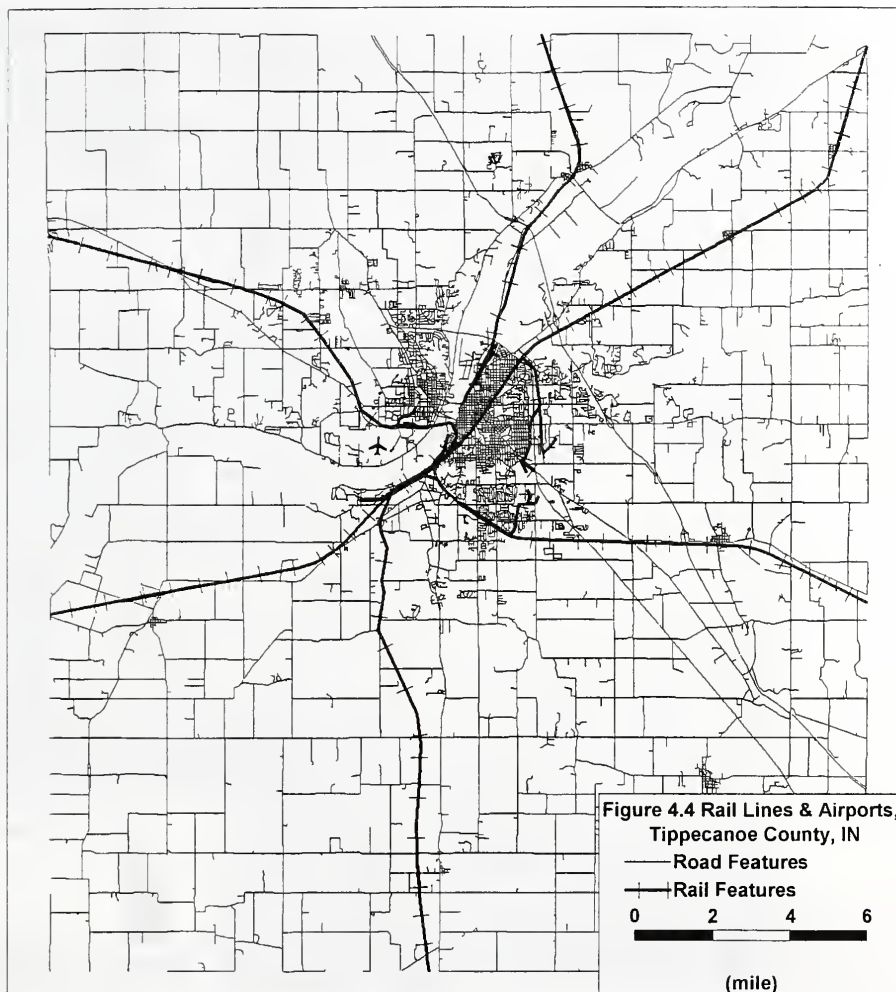
Both the coded network and trip matrices for all trip purposes are the inputs required for the traffic assignment stage.



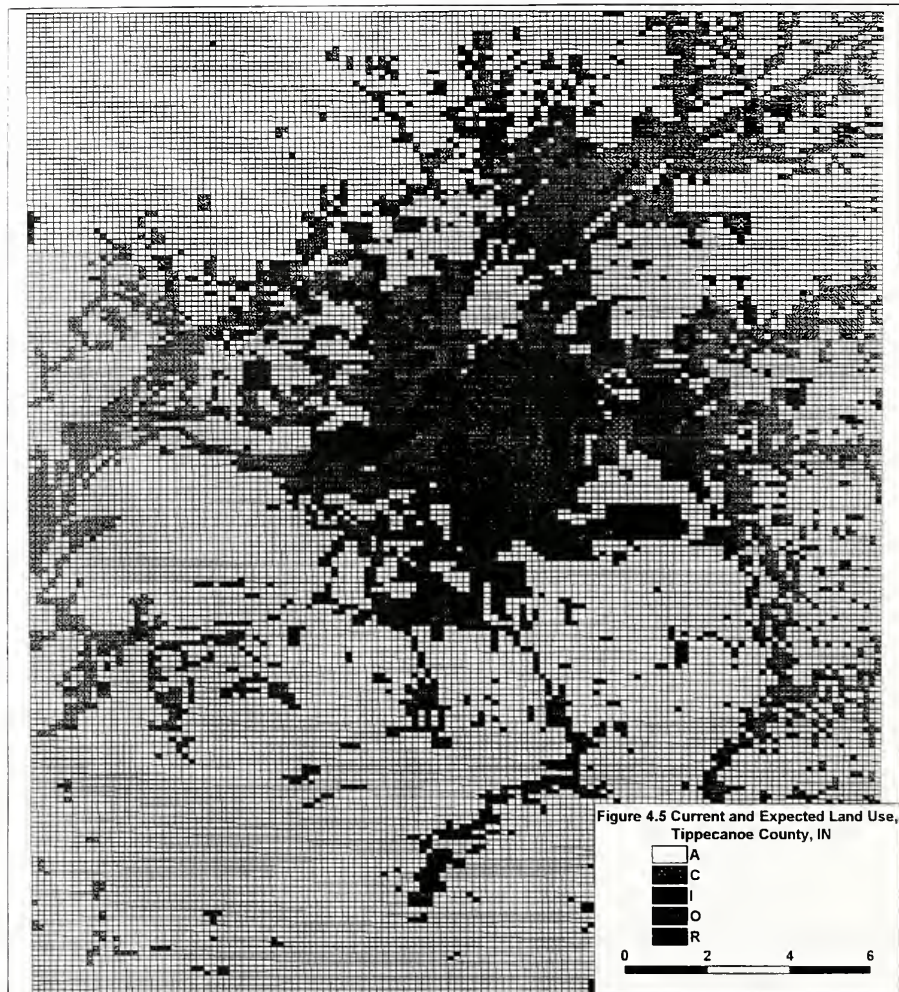














## CHAPTER 5 CASE STUDY

### 5.1 Tippecanoe County, Indiana

In order to apply the integrated model to the Lafayette area of Tippecanoe County, Indiana, the land within this study area is divided into three categories: urban, urbanizing and rural (Figure 5.1). There are 13 Townships in Tippecanoe County and 210 traffic analysis zones, including 182 internal zones and 28 external zones, in the Lafayette area to be analyzed. For the purpose of the following analysis, the Lafayette area is divided into 13 districts as follows (Figure 5.2).

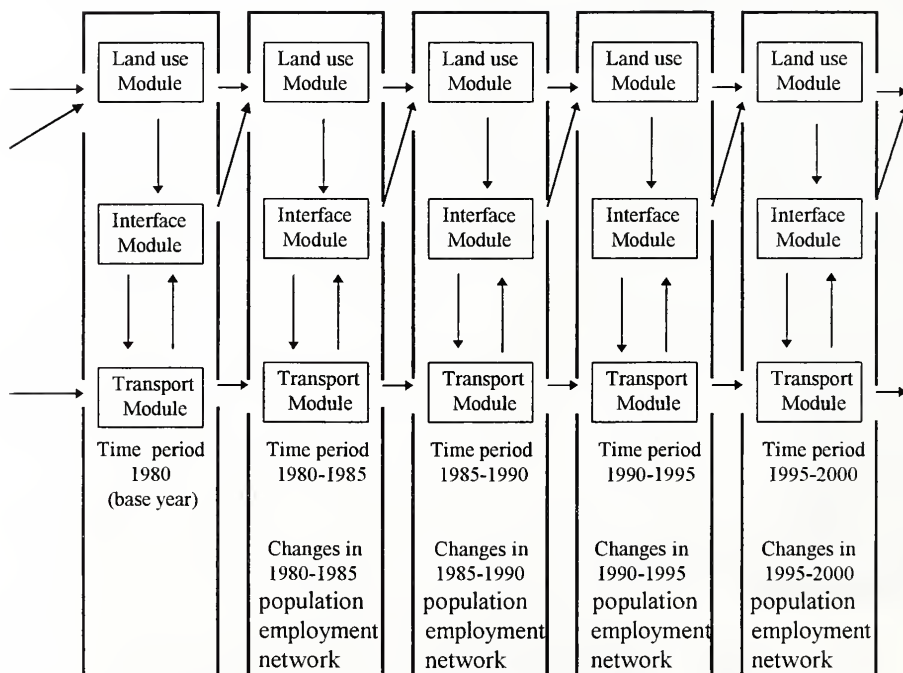
District 1: Shelby Township.	District 8: Wea Township.
District 2: Wabash Township.	District 9: Union Township
District 3: Tippecanoe Township.	District 10: Wayne Township.
District 4: Washington Township	District 11: Jackson Township.
District 5: Perry Township.	District 12: Randolph Township.
District 6: Sheffield Township.	District 13: Lauramie Township.
District 7: Fairfield Township.	

A brief recent statistical history of the Lafayette area is given in Table 5.1. The data are shown by district in Table 5.2. The 1985 data in Table 5.2 are computed by using linear interpolation between the 1980 and 1990 data.

**Table 5.1 Summary of the Lafayette area in 1970, 1980, 1990,1993, 2000**

	1970	1980	1990	1993	2000
Population	109,378	121,702	130,598	136,415	139,788
Housing unit	34,197	43,130	48,134	51,648	53,804
Total employment	52,250	63,439	70,703	72,339	77,833
Non-retail	43,110	51,555	56,667	58,620	61,956
Retail	9,140	11,884	14,036	13,719	15,877

A dynamic process of interaction between land use and transportation applied to the Lafayette area is shown on Figure 5.3. The integrated model is calibrated on 1980 (base year) data and changes over 1985-1990-1993. The application of the integrated model represents, together with forecasts for 2010 and 2015, the base case against which a set of policies can be tested. (See Sections 5.1.5, 5.1.6 and 5.1.7.)



**Figure 5.3 Order of model operation for simulating the dynamics of interaction between Land use and Transportation, Tippecanoe County, IN**

### 5.1.1 Land Use Allocation Module

The application of a land-use allocation model in the Lafayette area of Tippecanoe County, Indiana includes four major land-use activities: residential, nonretail industries, service industries for population, and service industries for businesses. Nonretail industries in the forecast year are assumed to be fixed at known locations. The model attempts to forecast the likely directions of population growth and retail employment in Tippecanoe County. Because of limitations on the data available, the model applied to the Lafayette area is used for one household category and two employment types, nonretail and retail employment. The residential location model and the employment location model applied to the Lafayette area are described below.

#### 1. The residential location model

The population of each zone in Lafayette area is computed from:

$$P_i = \sum_j \alpha E_j B_j' w_i' \exp(-\beta_1 t_{ij}) \quad (5.1)$$

and

$$B_j' = [\sum_i w_i' \exp(-\beta_1 t_{ij})]^{-1} \quad (5.2)$$

The term  $B_j' w_i' \exp(-\beta_1 t_{ij})$  represents the average probability of a commuter employed at zone j will choose to live in zone i.

where

$P_i$  : the population living in zone i at time period t.

$\alpha$  : the ratio of population to employees over the study area at time period t.

$E_j$  : the employees working in zone j at time period t.

$w_i'$  : the attractiveness of residential zone i at time period t.

$t_{ij}$  : the disutility of travel between zone i and zone j.

and,

$$w_i' = (L_i')^{\alpha_1} (DW_{i,r})^{\alpha_2} (R_i' + C_{i,j}')^{-\alpha_3} \quad (5.3)$$

and,

$$DW_{i,r} = S_i' * (U_{i,o}' + U_{i,r}') \quad (5.4)$$

where

- $L_i^t$  : net developable area for residential use in zone  $i$  at time period  $t$ .  
 $DW_{i,t}$  : the number of dwelling units offered for renter and owner occupancy in zone  $i$  at time period  $t$ .  
 $R_i^t$  : the average monthly rent of dwelling units in zone  $i$  during time period  $t$ .  
 $C_{i,j}^t$  : the average travel cost per month for employees who work in zone  $j$  and live in zone  $i$  during time period  $t$ .  
 $S_i^t$  : the total number of housing units in zone  $i$  at time period  $t$ , including vacant housing. This value can be calculated using equation (5.21).  
 $U_{i,r}^t$  : the probability that a housing unit in zone  $i$  at time period  $t$  will be offered for renter occupancy, given the ongoing average rent  $R_i^t$ . This value can be calculated by equation (5.19).  
 $U_{i,o}^t$  : the probability that a housing unit of household type  $n$  in zone  $i$  at time period  $t$  will be offered for owner occupancy, given the ongoing average rent  $R_i^t$ . This value is obtained by equation (5.18).  
 $\beta_i, \alpha_1, \alpha_2, \alpha_3$ : empirically calibrated parameters.

Adjustments to residential attractiveness

If a zone is almost fully occupied by retail or nonretail activities (or both), then only a few people should be able to live there. The model can be instructed to reduce residential attractiveness in response to a large allocation of retail or nonretail employees (or both) to a zone. This adjustment is handled iteratively. The residential developable area in zone  $i$  at time period  $t$  is based on the number of retail and nonretail employees allocated to zone  $i$  at the same time period and the amount of new dwellings constructed and old dwellings demolished during time period  $t-1$ .

$$L_i^t = L_i - z b_i * E b_i^t - z s_i * E s_i^t + Z_i^{t-1} * (D_i^{t-1} - C_i^{-1}) \quad (5.5)$$

where

- $L_i^t$  : the residential developable area in zone  $i$  at time period  $t$ .  
 $L_i$  : net developable area for residential and employment development in zone  $i$ .  
 $z b_i$  : land area occupied per nonretail employee in zone  $i$ .  
 $E b_i^t$  : the amount of employees in nonretail employment in zone  $i$  at time period  $t$ .  
 $z s_i$  : land area occupied per retail employee in zone  $i$ .  
 $E s_i^t$  : the amount of employees that serve population and employees in zone  $i$  at time period  $t$ .  
 $Z_i^{t-1}$  : an average density coefficient for dwelling units in zone  $i$  during time period  $t-1$ .

( $Z_i^{n,t-1}$  can be calculated based on the criteria listed in Table 3-1)

$C_i^{t-1}$  : the number of new dwelling units constructed in zone i at time period t-1.

$D_i^{t-1}$  : the number of old dwelling units demolished in zone i at time period t-1.

## 2. The employment location model

The retail employment for each zone in the Lafayette area is calculated from:

$$E_j^{r,t} = \sum_i \delta (P_i^t + E_i^{nr,t} + E_i^{r,t}) \bar{B}_i^t S_j \exp(-\beta_2 t_{ij}) \quad (5.6)$$

where

$$\bar{B}_i^t = [\sum_j S_j \exp(-\beta_2 t_{ij})]^{-1} \quad (5.7)$$

The term  $\bar{B}_i^t S_j \exp(-\beta_2 t_{ij})$  represents the average probability of a person living (or working) in zone j will choose zone j for shopping or other retail activities.

where

$E_j^{r,t}$  : the number of retail employees working in zone j at time period t.

$\delta$  : the ratio of the number of retail employees to the sum of population and employees across the whole study area at time period t.

$P_i^t$  : the population living in the zone i at time period t.

$E_i^{nr,t}$  : the number of nonretail employees working in zone i at time period t.

$E_i^{r,t}$  : the number of retail employees working in zone i at time period t.

$t_{ij}$  : the disutility of travel between zone i and j.

and,

$$S_j = (SL_j^t)^{\gamma_1} (SE_j^{t-1} / L_j^{T,t})^{\gamma_2} \quad (5.8)$$

where

$S_j$  : the attractiveness for retail employment in zone j at time period t.

$SL_j^t$  : developable area for retail (or service) activity in zone j at time period t.

$SE_j^{t-1}$  : the number of retail employees in zone j at time period t-1.

$L_j^{T,t}$  : total land area in zone j at time period t.

$\beta_2, \gamma_1, \gamma_2$ : empirically calibrated parameters

Adjustments to service attractiveness

An urban agglomeration is a geographic concentration of economic activity. Agglomeration is a notion that spatial proximity of activities makes resources more efficient than if such activities are spatially dispersed (Goldstein and Gronberg, 1984). Even in a small urban area, significant agglomeration in the service sector can occur. Agglomeration will occur in districts that have the highest potential to attract a disproportionately large share of suppliers (or developers). These districts will tend to be central to the region or near large concentrations of population. In the employment location program, the measure of potential is continually updated from one time period to the next.

$$S_j^{t+1} = (SL_j^{t+1})^{\gamma_1} (SE_j / L_j^{T,t+1})^{\gamma_2} \quad (5.9)$$

where

$S_j^{t+1}$  : a measure of service (or retail) attractiveness in zone j at time period t+1.

$SL_j^{t+1}$  : a measure of service (or retail) attractiveness (developable service area) in zone j at time period t+1.

$SE_j^t$  : the number of retail employees in zone j at time period t.

$L_j^{T,t+1}$  : total land area in zone j at time period t+1.

$\gamma_1$  : calibrated parameter for service attractiveness.

$\gamma_2$  : central tendency parameter; controls how much agglomeration (i.e., the tendency for businesses to locate near each other) will occur in the service sector.

#### Calibration of land use allocation model

In the residential location model, the calibration of parameters (e.g.,  $\beta_1$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_3$ ) in equations (5.1)-(5.3) must satisfy:

$$\text{Max } \psi_{pop} = \sum_i P_i \ln \hat{P}_i - \sum_i P_i \ln P_i \quad (5.10)$$

$$\text{subject to } \bar{C}_{\min} p \leq \bar{C}_{\text{mod}} p \leq \bar{C}_{\max} p \quad (5.11)$$

where

$P_i$  : the observed population living in the zone i.

$\hat{P}_i$  : the estimated population living in the zone i by the integrated model.

$\bar{C}_{\max}^p$  : the upper limit of mean work-to-home trip cost.  
 $\bar{C}_{\text{mod}}^p$  : the estimated mean work-to-home trip cost.  
 $\bar{C}_{\min}^p$  : the lower limit of mean work-to-home trip cost.

And, in the employment location model, the calibration of parameters (e.g.,  $\beta_2$ ,  $\gamma_1$ ,  $\gamma_2$ ) in equations (5.6)-(5.8) must satisfy:

$$\text{Max } \psi_{\text{ret}} = \sum_j E_j \ln \hat{E}_j - \sum_j E_j \ln E_j \quad (5.12)$$

$$\text{subject to } \bar{C}_{\min}^e \leq \bar{C}_{\text{mod}}^e \leq \bar{C}_{\max}^e \quad (5.13)$$

where

$E_j$  : the observed employees working in zone j.  
 $\hat{E}_j$  : the estimated employees working in zone j by the integrated model.  
 $\bar{C}_{\max}^e$  : the upper limit of mean home-to-shop trip cost.  
 $\bar{C}_{\text{mod}}^e$  : the estimated mean home-to-shop trip cost.  
 $\bar{C}_{\min}^e$  : the lower limit of mean home-to-shop trip cost.

The estimated values of these parameters are listed in Table 5.3. All these parameters are statistically significant at the 5% level, except the parameters  $\beta_1$  and  $\beta_2$ . Because  $\beta_1$  and  $\beta_2$  must satisfy the constraints (5.11) and (5.13) respectively, it is not necessary to calculate their asymptotic t-statistic.

The term  $\exp(-\beta_1 t_{ij})$  in equation (5.2) is a measure of an employee's willingness to make long trips between his/her workplace and home. It is expected that  $\beta_1$  will be positive. This means that the longer the home-based work trip is, the less willing an employee will be to make it. At the same time, the calibrated value of  $\beta_1$  must ensure that the model's mean home-based work trip cost is equal to the observed mean home-based work trip cost. Furthermore, the term  $\exp(-\beta_2 t_{ij})$  in equation (5.7) is a measure of peoples' willingness to make long trips between home and shopping place. In a way similar to  $\beta_1$ ,  $\beta_2$  is expected to be positive and must ensure that the model's mean home-based shopping trip cost is equal to the observed mean home-based shopping trip cost. In



Table 5.3,  $\beta_1$  and  $\beta_2$  have the expected signs when they are calibrated on 1980, 1990 and 1993 data.

The  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  exponents in equation (5.3) are for developable area for residential use, the number of dwelling units, and the sum of monthly rent and travel cost, respectively. In general, if a zone has more area available for residential use, more population will be able to live there. Therefore,  $\alpha_1$  is expected to be positive. Similarly, if a zone has more dwelling units offered for renter and owner occupancy, then it can accommodate more population living there. So  $\alpha_2$  is expected to have a positive sign. In equation (5.3), the term  $(R_{i,j} + C_{i,j}^t)$  is the sum of rent and travel cost that will be paid by employees each month. It is assumed that employees will choose their house location by trading off the monthly rent of the house versus the travel cost they will pay each month. Because employees try to pay as little as possible for the combined expenses of housing and travel costs.  $\alpha_3$  is expected to be positive. In Table 5.3, the signs of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$  are positive, as expected.

The  $\gamma_1$  and  $\gamma_2$  in equation (5.8) are the exponents for developable area for retail activity and density of retail employment for each zone, respectively. If a zone has more developable area for retail use, it can accommodate more retail employees working there. Therefore,  $\gamma_1$  is expected to be positive. The term  $(SE_j^r / L_j^{T,j+1})$  represents the agglomeration that often occurs in the service (or retail) sector. The higher the density of retail employment in a district is, the more attractiveness for retail activity will be generated there. Thus,  $\gamma_2$  is expected to have a positive sign. Table 5.3 shows that the signs of  $\gamma_1$  and  $\gamma_2$  are positive, as expected.

In Table 5.3, the values of  $\psi_{pop}$ ,  $RMSE_{pop}$  and  $PRMSE_{pop}$  show that the land use allocation module can achieve a good fit to observed population data. The values of  $\psi_{ret}$  and  $PRMSE_{ret}$  show that the land use allocation module may generate a larger error for



retail employment, but these errors are still in an acceptable range. In the residential location model,

$$\text{RMSE}_{pop} = \sqrt{(1/N) \sum_{i=1}^N (P_i - \hat{P}_i)^2} \quad (5.14)$$

$$\text{PRMSE}_{pop} = \text{RMSE}_{pop} / (\sum_i P_i / N) \quad (5.15)$$

where

$\text{RMSE}_{pop}$  : root-mean-square error for population.

$\text{PRMSE}_{pop}$  : percent root-mean-square error for population.

$P_i$  : observed population living in the zone  $i$ .

$\hat{P}_i$  : estimated population living in the zone  $i$  by the integrated model.

$N$  : the number of zones in the study area.

In the employment location model,

$$\text{RMSE}_{ret} = \sqrt{(1/N) \sum_{j=1}^N (E_j - \hat{E}_j)^2} \quad (5.16)$$

$$\text{PRMSE}_{ret} = \text{RMSE}_{ret} / (\sum_j E_j / N) \quad (5.17)$$

where

$\text{RMSE}_{eml}$  : root-mean-square error for retail employment.

$\text{PRMSE}_{ret}$  : percent root-mean-square error for retail employment.

$E_j$  : observed employees working in the zone  $j$ .

$\hat{E}_j$  : estimated employees working in the zone  $j$  by the integrated model.

$N$  : the number of zones in the study area.

In Table 5.3, two observations can be made. First, the calibration to the 1980 data does not indicate as good a fit as do the 1990 and 1993 calibrations. Second, the parameter values and goodness-of-fit measures for 1990 and 1993 are in fairly close agreement. Although there is only a 3-year gap between these calibrations, the calibrations are based on two separate data files that include some noticeable shift in population and employment (see Table 5.2). The similar results for 1990 and 1993 seem to indicate that the model is not overly sensitive to moderate changes in input. Because the 1990 and 1993 calibration

results confirm each other, the set of calibrated parameters for 1993 (the most recent year) in Table 5.3 will be used to simulate the spatial distributions of population and retail employment for 2010 and 2015. (See Sections 5.1.5 - 5.1.7.)

### **5.1.2 Land Use Potential Study**

As mentioned in Section 3.3, the work in the land use potential study can be used as the input for computing developable area for specific activities (e.g., residential, commercial or other development). It, moreover, has the following capabilities (APC, 1981):

- 1) The role of public utilities and services as basic elements that influence urban development can be assessed.
- 2) Strategies to protect prime agricultural land, open space and natural areas by diverting development to more suitable areas can be devised and tested.
- 3) The way of industrial and commercial growth that can be expected to respond to transportation network and utility expansion can be investigated. For example, industrial and commercial use potential is made highest at the intersection of major roadways rather than between these intersections.

The results of a land use potential study -- especially one that is automated -- can be easily adjusted as factors change. Thus the construction or extension of a sanitary sewer line can be accompanied by a systematic review of the land use plan. New land use potential scores can be obtained by applying the point system method (APC, 1981) and Dempster-Shafer theory to update land use potential maps. Any subsequent changes in the current land use plan and highway network can be justified and documented. In a similar way, updated soil information can be used to revise the land use potential study.

In order to provide a useful tool to perform the job of allocating sufficient land to meet the needs of development in the Lafayette area, APC (1981) proposed a point system approach (see Appendix A) to make rational and consistent decisions about how land ought to be used in the future. Based on the results of the point system method described earlier, APC defined a land use plan as three stages. They are:

- 1) Current and expected land use plan: how land is being used at the present time, including development plans that are working their way through the approval process (Figure 5.4).
- 2) First cut land use plan: how land would be used – regardless of actual need -- if it were to be used to its highest potential.
- 3) Phased land use plan: how land ought to be used to meet the needs of development, including residence, farming and recreation, in the foreseeable future.

As mentioned in Section 3.2, the “Current and expected land use” is one of the eight equally weighted factors used by the point system approach. The first cut land use plan is based entirely on the results of the point system procedure. Figure 5.5 shows an expanded version of the current and expected land use plan for the Lafayette area that could be used to its best potential (APC, 1981).

In general, the ultimate development potential obtained from the first cut land use plan is well beyond its actual needs for the foreseeable future. Therefore, a methodology to systematically and objectively convert the first cut land use plan into “Phased land use plan” would greatly assist the land use planning effort. (APC, 1981) The methodology could be used to pare down the amount of land with strong potential for long-term residential development to a level more appropriate to Tippecanoe County’s shorter-term population and economic growth forecasts. APC (1981) used a residential land use potential study to accomplish this objective. In fact, the result obtained from Dempster-Shafer theory in this research could have been used as the basis for APC’s actual “Phased land use plan”. A map produced from Dempster-Shafer theory is shown in Figure 5.6, and it closely resembles APC’s phase land use plan map. A map such as this can readily serve as a database from which a practical land use map and plan can be derived, incorporating the intuition and experience of planners, the comments of the general public who are asked to review the plan, and the feelings of public officials who are asked to adopt the plan.

### 5.1.3 Land Consumption Model

In the Lafayette area, the probabilities  $U'_{i,o}$  and  $U'_{i,r}$  that an existing housing unit will be offered for owner occupancy and renter occupancy can be derived using the following equations.

$$U'_{i,o} = \eta_{o,0} + \eta_{o,1} R'_i + \eta_{o,2} X'_{i,1} + \mathcal{E}'_{i,o} \quad (5.18)$$

$$U'_{i,r} = \eta_{r,0} + \eta_{r,1} R'_i + \eta_{r,2} X'_{i,1} + \mathcal{E}'_{i,r} \quad (5.19)$$

and satisfy

$$U'_{i,o} + U'_{i,r} + U'_{i,v} = 1.0 \quad (5.20)$$

where

$U'_{i,v}$  : the probability that a housing unit in zone i will remain vacant at time period t.

$R'_i$  : the average monthly rent of housing units in zone i during time period t.

$X'_{i,1}$  : location dummy variable; if zone i is located in Lafayette city or West Lafayette city, then  $X'_{i,1} = 1$ , otherwise  $X'_{i,1} = 0$ .

$\eta_{o,0}, \eta_{o,1}, \eta_{o,2}, \eta_{r,0}, \eta_{r,1}$  and  $\eta_{r,2}$  are empirically estimated coefficients.

$\mathcal{E}'_{i,o}$  and  $\mathcal{E}'_{i,r}$  are random error terms.

The estimates for coefficients in equations (5.18) and (5.19) are based on the 1990 census data. The average monthly rents for townships in Tippecanoe County are given in Table 5.4. The values of the estimated coefficients are listed in Table 5.5. These coefficients are statistically significant at the 5% level, except the coefficient for monthly rent ( $R'_i$ ) for owner occupancy is statistically significant at the 10% level.

Because only data on the net change,  $ND_i^{-1}$ , in housing units for each year are available to APC, the housing stock adjustment submodel applied to the Lafayette area adjusts the number of housing units for each zone using the following formula:

$$S_i^t = S_i^{t-1} + ND_i^{t-1} \quad (5.21)$$

and

$$ND_i^{t-1} = C_i^{t-1} - D_i^{t-1} \quad (5.22)$$

where

$S_i^t$  : the number of housing units in zone i at time period t.

$C_i^{t-1}$  : the number of new housing units constructed in zone i during time period t-1.

$D_i^{t-1}$  : the number of old housing units demolished in zone i during time period t-1.

But the data of  $C_i^{t-1}$  and  $D_i^{t-1}$  are not available in the Lafayette area. Thus,

$$DW_{i,t} = S_i^t * (U_{i,o}^t + U_{i,r}^t) \quad (5.23)$$

where

$DW_{i,t}$  : the number of dwelling units offered for renter and owner occupancy in zone i at time period t.

#### 5.1.4 Travel Demand Model

##### 1) Trip generation

In the trip generation stage, the trip productions for HBO and NHB at each time period are derived using the following equations (TETUP, 1994).

$$\text{HBO trip productions for each zone} = P_{HBO,i} * (DW_i * TRIP_{i,DW}) \quad (5.24)$$

$$\text{NHB trip productions for each zone} = P_{NHB,i} * (DW_i * TRIP_{i,DW}) \quad (5.25)$$

$$\text{E-I trip productions} = \text{E-I trip attractions (see equation (5.32))} \quad (5.26)$$

Also,

$$(\text{Popgen})_i = 4.77048 * [(\text{Pop})_i / DW_i] \quad (5.27)$$

$$(\text{Autogen})_i = 6.0858 * [(\text{Auto})_i / DW_i] \quad (5.28)$$

$$TRIP_{i,DW} = 0.5 * [0.30804 + (\text{Popgen})_i + (\text{Autogen})_i] \quad (5.29)$$

where

$P_{HBO,i}$  : the percent of average daily person trips produced for HBO purpose in zone i.

- $P_{NHBj}$  : the percent of average daily person trips produced for NHB purpose in zone i.
- $DW_i$  : the number of dwelling units in zone i.
- $TRIP_{i,DW}$  : the trip production rate of each dwelling unit in zone i.
- $(Popgen)_i$  : the number of trips generated by each household in zone i.
- $(Autogen)_i$  : the number of trips generated by each household based on auto ownership in zone i.
- $(Pop)_i$  : the population in zone i.
- $(Auto)_i$  : the amount of vehicle in zone i.

$P_{HBOj}$  and  $P_{NHBj}$  can be chosen depending on the amount of population in each time period for the Lafayette area.

	$P_{HBOj}$	$P_{NHBj}$
Population $\leq$ 100,000	0.3929	0.3927
100,000 < Population $\leq$ 250,000	0.3529	0.3927

Source: Tippecanoe County Area Plan Commission, Indiana, 1994.

The trip attractions for HBO and NHB in each time period are calculated using the following equations (TETUP, 1994).

$$\text{HBO trip attractions for each zone} = 10.0 * R_i + 0.5 * NR_i + 1.0 * DW_i \quad (5.30)$$

$$\text{NHB trip attractions for each zone} = 2.0 * R_i + 2.5 * NR_i + 0.5 * DW_i \quad (5.31)$$

$$\text{E-I trip attractions} = 67.0 + 2.28 * R_i + 0.5 * NR_i + 0.2 * DW_i \quad (5.32)$$

where

- $R_i$  : the number of retail employees in zone i.
- $NR_i$  : the number of nonretail employees in zone i.
- $DW_i$  : the amount of dwelling units in zone i.

In the Lafayette area, the regional total for trip productions, by purpose, is used as the control trip total. Because decennial census data are used to identify the household characteristics by zone, we have a greater degree of confidence in household data than we do in employment data. Thus zonal trip attractions will be balanced to match regional productions.

## 2) Trip distribution

The distribution of HBW trips in the Lafayette area is accomplished in the land-use allocation module (see Section 5.2). The distributions of HBO, NHB and external-internal or internal-external trips are performed by using a gravity model in a commercial package, TRANPLAN. The external-external trips are estimated based on 1980 survey data from APC.

## 3) Modal split

In the Lafayette area, the percentage of transit person trips is less than 2 percent. The transit person trips between each zone pair are converted into vehicle trips by using a bus occupancy rate (e.g., 5 persons per bus). Likewise, the auto person trips are converted into vehicle trips using a set of auto occupancy rates. (See Table 3.2.) Then, the total vehicle trips between each zone pair will be loaded on the network in the traffic assignment stage.

## 4) Traffic assignment

In the trip distribution stage, the vehicle trips for each purpose are combined into a trip matrix between each zone pair. The assignment process is performed by using an equilibrium assignment approach under capacity restraint, and with turn prohibitors and turn penalties. The assignment procedure is iterated until convergence, that is, until there is no change in travel time (or transport costs) from one iteration to the next. The generating disutilities (time or costs) of the traffic assignment stage are fed back into the



land use allocation module, thus influencing the location of activities (residence and retail development) in the next time period. (See Figure 5.3.)

### **5.1.5 Policy Test 1 - improvements of highway network for 2010**

The first policy test is an analysis to assess the effect of changes in the highway network in the Lafayette area for the year 2010. A realignment of US 231 will cross the Wabash River from South River Road to County Road 350S, then connect to the current US 231 south of McCutcheon High School near County Road 500S. Furthermore, County Road 350S will be expanded to a major two-lane highway connecting State Road 38 (at County Road 475E) west to the new US 231. This creates a southern bypass around the Lafayette area from I-65 to Purdue University. (See Figure 5.7.) All these changes in the highway network are scheduled to be accomplished by the year 2000.

The first policy test is run by using the integrated Land-use/Transportation model that has been described in Chapter 3. The resulting population forecasts and retail employee forecasts in the year 2010 are shown in Table 5.6 and Table 5.8, respectively. The “base run” listed in Tables 5.6 and 5.8 represents a year 2010 base case (i.e., no realignment of US 231). Then, Policy Test 1 is carried out by using the same set of parameters and monthly rent for each zone as those used in the base run, but with a realigned US 231 included. Policy Test 1 is iterated until convergence, that is, the integrated model reaches a stabilized situation between the land use allocation module and the travel demand module during the same time period. (See Figure 3.1.) A stabilized situation means that there is no significant change in the distribution of population and retail employees over the study area in the land use allocation module, and no significant change in travel time (or transport costs) in the travel demand module from one iteration to the next. As Tables 5.6 and 5.8 indicate Policy Test 1 reached a stabilized situation after only three iterations. The differences in population between the base run, Policy Test 1, and forecasts by APC are shown in Table 5.7. (The APC forecasts do not include a realigned US 231 when APC made a future land use forecasts.) The differences in retail employees between the base run, Policy Test 1, and forecasts by APC are shown in Table 5.9.



Table 5.7 shows that there are significant differences in the distribution of population between base case and APC forecasts for the year 2010. The reason is that the APC forecasts did not simultaneously consider the changes in the road network in the Lafayette area when APC made a future land use forecasts for the year 2010. One significant change is County Road 350 S expanded to be a continuous major two-lane highway connecting SR 38 (at County Road 475 E) west to the new US 231. Other changes are the following roads that are planned to be improved to four lanes (see Figure 5.8):

- 1) Lindberg Road from Northwestern Avenue to McCormick Road.
- 2) Creasy Lane from Greenbush Street to SR 38.
- 3) McCarty Lane from Main Street to Creasy Lane.
- 4) Union Street from 21<sup>th</sup> Street to Creasy Lane.
- 5) South River Road from SR 26 to the north end of realignment of US 231.

These changes in the network are scheduled to be completed by the year 2000 and all of them are considered in the base case for the year 2010.

All above changes in the road network are located in Fairfield Township (District 7), Sheffield Township (District 6), and Wabash Township (District 2). These changes will improve the traffic situation in these three townships. Thus, these three townships will become more accessible due to these changes by 2010. Therefore, more people will be willing to shift to or stay in Fairfield Township (District 7), Sheffield Township (District 6), and Wabash Township (District 2) in the base case than those forecasted by APC.

A major benefit from the realignment of US 231 is that it will significantly improve accessibility between the West Lafayette area and southern Tippecanoe County (Figure 5.7). Therefore, Tables 5.6 and 5.7 show a spatial distribution of population for 2010 based on Policy Test 1 that has some reasonable characteristics. The five northern townships, including Lafayette city and West Lafayette city, will lose a small amount of population. Most of this population shift is into the eight southern townships. The largest population increases will be in southern Fairfield Township, Wea Township (District 8), Union Township (District 9) and Sheffield Township (District 6) near the south end of the

newly constructed road. This is to be expected, because new US 231 will become a major road connecting Wabash Township, Purdue University (District 2) and Wea Township (District 8) and Randolph Township (District 12). Furthermore, the congested condition on old US 231 will be improved due to a large amount of traffic shifting to the new US 231. Thus, travel from Fairfield Township (District 7) to Wea Township (District 8) and to Randolph Township (District 12) will become more accessible because of the changes in US 231. (See Table 5.10.) The values of V/C in the old US 231 for 1993 and 2010 are listed in Table 5.11.

In Table 5.8 and Table 5.9, the results of Policy Test 1 show that significant retail employment increases will occur in southern Fairfield Township (District 7) near the south end of the realignment of US 231. One reason is that the Fairfield Township has a large amount of developable commercial area, including Tippecanoe Mall, available for retail employment in 2010. Another reason is that the retail employment in the forecast year will change in direct proportion to changes in the distribution of population (or households). This implies that more population (or households) clustered in an area that has better accessibility to workplace will generate more retail employment to serve this additional population (or households). (See Table 5.6.)

#### **5.1.6 Policy Test 2 - changes in land use pattern**

This policy test is an analysis to assess the effect of changes in the land use pattern in the Lafayette area. The policy test indicates that the 1981 adopted Comprehensive Plan for Tippecanoe County called for little residential expansion in the mostly rural sector between I-65 on the west, Wildcat Creek on the north and east, and the Town of Dayton on the south. (See Figure 5.9.) Expansion of utilities along SR 26, and along Haggerty Lane (as a result of the SIA plant locating just west of Dayton), have changed the development equation. All these changes in the utilities are scheduled to be accomplished by the year 2000. A considerable amount of residential growth is anticipated in these traffic zones. It is necessary to determine the roadway network necessary to handle the

sharply increasing traffic that will result. Policy Test 2 also considers the realignment of US 231.

The second Policy Test is accomplished by using the integrated model. The resulting population forecasts and retail employee forecasts in the year 2010 are shown in Tables 5.12 and 5.13, respectively. Policy Test 2 is carried out by using the same set of parameters and monthly rent for each zone as those used in Policy Test 1. Policy Test 2 is iterated until reaching a stabilized situation between the land use allocation module and the travel demand module from one iteration to the next during the same period. (See Table 5.12.) As Tables 5.12 and 5.13 indicate Policy Test 2 reached a stabilized situation after only two iterations. The differences in population between Policy Test 2 and forecasts by APC are shown in Table 5.12. (The APC forecast for 2010 also considers the residential development in the shaded area in Figure 5.9.) The differences in retail employees between Policy Test 2 and forecasts by APC are shown in Table 5.13.

The expansion of utilities to the shaded area (see Figure 5.9) will create significant residential growth in this area after the year 2000. As a result of Policy Test 2, a little residential expansion is expected to appear in the shaded area (southern Perry Township and northern Sheffield Township) for the year 2010. (See Table 5.12.)

One result produced by Policy Test 2 is worth noting. The traffic conditions on SR 26 will get worse by the year 2010 (see Table 5.14). It is necessary to make an improvement on SR 26 between US 52 and County Rd. 700E by the year 2010, especially the section between I-65 and County Rd. 700 E on SR 26.

### **5.1.7 Policy Test 3 - improvements of highway network for 2015**

The third policy test is an analysis to assess the effect of particular proposed changes in the highway network on the West Lafayette area for the year 2015. The realignment of US 231 in Policy Test 1 is extended here to create a western bypass around the West Lafayette area. A new four-lane road will carry traffic from the US 231/South River Road

intersection around the east and north sides of Purdue Airport, across State Road 26 on up to McCormick Road, just before it crosses Lindberg Road. From there, McCormick Road will be widened to four lanes up to a new two-lane, southwest-to-northeast connector to Cumberland Avenue at US 52. The north end of McCormick will become a four-lane facility. (See Figure 5.10.) All these changes in the highway network are scheduled to be accomplished by the year 2010.

The third policy test is accomplished by the integrated Land-use/Transportation model. The resulting population forecasts and retail employee forecasts in the year 2015 are shown in Table 5.15 and Table 5.17, respectively. The “base run” listed in Tables 5.15 and 5.17 represents a year 2015 base case (i.e., no realignment of US 231 north of River Road in the West Lafayette area). Then, Policy Test 3 is carried out by using the same set of parameters and monthly rent for each zone as those used in the base run, but with a realigned US 231 north of River Road included. As Tables 5.15 and 5.17 indicate Policy Test 3 reached a stabilized situation after only two iterations. The differences in population between the base run, Policy Test 3, and forecasts by APC are shown in Table 5.16. (The APC forecasts do not include any realignment of US 231, even south of River Road when APC made a future land use forecasts.) The differences in retail employees between the base run, Policy Test 3, and forecasts by APC are shown in Table 5.18.

Table 5.16 shows that there are significant differences in the distribution of population between base case and APC forecasts for the year 2015. The reason is that the APC forecasts for the year 2015 did not consider the changes in the Lafayette area road network. For example, APC did not consider the realignment of US 231 in the southern Lafayette area (i.e., Policy Test 1) when they made land use forecasts for the year 2015. APC also did not consider the following changes (see Figure 5.11):

- 1) South Street will be reconfigured to three lanes from 18<sup>th</sup> Street to Earl Avenue.
- 2) A new four-lane, Park Drive East, will connect State Road 26 with McCarty Lane, just west of I-65.

The following roads are planned to be improved to four lanes:

- 3) Greenbush Street from US 52 to Creasy Lane.
- 4) Earl Avenue from Union Street to Main Street.
- 5) Teal Road from 4<sup>th</sup> Street to 18<sup>th</sup> Street.
- 6) Brady Lane from 18<sup>th</sup> to US 52.
- 7) South 4<sup>th</sup> Street from Teal Road to State Road 25.
- 8) Pre-alignment US 231 from State Road 25 to Beck Lane, and
- 9) State Road 25 from Workman Lane to pre-alignment US 231.

These changes in the network are scheduled to be completed by the year 2010 and all of them are considered in the base case for the year 2015.

All of the above changes in the road network are located in Fairfield Township (District 7), Sheffield Township (District 6) and Wea Township (District 8). These changes will improve the traffic situation in these three townships, making them more accessible to/from other key trip end locations. As a result, more people will be willing to shift to or stay in Fairfield Township (District 7), Sheffield Township (District 6), and Wea Township (District 8) in the base case than those forecasted by APC for the year 2015.

A major benefit from the realignment of US 231 in the West Lafayette area is that it will significantly improve accessibility to/from West Lafayette. In particular, travel between the West Lafayette area and southern Tippecanoe County will become easier because of the changes in US 231. (See Table 5.19.) Therefore, from the results of Table 5.15 and Table 5.6, two observations can be made. First, they show that little difference between Policy Test 1 and 3 in Wabash Township (District 2) population for the years 2010 and 2015. Second, the spatial distribution of population for 2015 based on Policy Test 3 is similar to that generated by Policy Test 1. (See Section 5.1.5.) Furthermore, the congestion on Northwestern Avenue and US 52 through West Lafayette will be alleviated by traffic shifting to the new northern realignment of US 231. The values of V/C on Northwestern Avenue and US 52 for 1993 and 2015 are listed in Table 5.20.

In Table 5.17 and Table 5.18, the results of Policy Test 3 also show that significant retail employment increases will occur in southern Fairfield Township (District 7), as was

the case in Policy Test 1. The reason is that the Fairfield Township still has a large amount of developable commercial area, including Tippecanoe Mall, available for retail employment in 2015. This situation shows that agglomeration will occur in districts that have the highest potential to attract a disproportionately large share of suppliers (or developers).

**Table 5.2 Summary of the Lafayette area by districts in 1980, 1985, 1990, 1993**

	District	Population	Retail Employment	Nonretail Employment	Total Employment
1980	1	1,855	16	2,079	2,095
	2	44,267	3,266	14,778	18,044
	3	4,636	35	1,197	1,232
	4	2,394	18	526	544
	5	2,720	1	428	429
	6	2,254	48	1,026	1,074
	7	44,533	8,086	2,4161	32,247
	8	12,698	290	2,987	3,277
	9	1,713	28	1,353	1,381
	10	1,233	12	678	690
	11	520	4	620	624
	12	754	20	528	548
	13	2,125	60	1,194	1,254
	Total	121,702	11,884	51,555	63,439
1985	1	1,927	18	1,269	1,287
	2	46,320	3,684	18,653	22,336
	3	4,682	58	917	975
	4	2,327	23	341	364
	5	2,779	2	328	330
	6	2,591	60	1,538	1,598
	7	45,834	8,727	25,072	33,799
	8	13,426	280	2,724	3,004
	9	1,543	17	1,403	1,420
	10	1,264	12	402	414
	11	519	4	412	416
	12	744	20	334	354
	13	2,051	55	697	752
	Total	126,007	12,960	54,090	67,049
1990	1	2,002	21	461	482
	2	49,348	4,100	22,528	26,628
	3	5,031	80	641	721
	4	2,346	29	158	187
	5	3,030	3	229	232
	6	2,414	71	2,055	2,126
	7	46,166	9,371	26,002	35,373
	8	14,066	270	2,466	2,736
	9	1,674	5	1,455	1,460
	10	1,196	12	126	138
	11	515	4	204	208
	12	628	20	141	161
	13	2,182	50	201	251
	Total	130,598	14,036	56,667	70,703



1993	1	2,010	21	461	482
	2	50,045	3,350	23,633	26,983
	3	5,229	80	691	771
	4	1,983	29	183	212
	5	3,717	3	379	382
	6	3,567	71	2,580	2,651
	7	48,178	9,703	25,355	35,058
	8	15,985	371	2,436	2,807
	9	1,378	5	2,230	2,235
	10	1,014	12	126	138
	11	753	4	204	208
	12	1,136	20	141	161
	13	1,420	50	201	251
	Total	136,415	13,719	58,620	72,339

**Table 5.3 Calibrated parameters of the land use allocation module on 1980, 1990 and 1993 data**

parameters	1980	1990	1993
	Asymptotic t-statistics	Asymptotic t-statistics	Asymptotic t-statistics
$\beta_1$	0.1810	0.1332	0.1076
$\beta_2$	0.2973	0.2337	0.2015
$\alpha_1$	1.1101 (5.60)	0.8099 (6.01)	0.7119 (6.30)
$\alpha_2$	0.3466 (2.92)	0.5208 (3.81)	0.4363 (3.92)
$\alpha_3$	0.8821 (11.65)	1.7017 (13.88)	1.6247 (13.62)
$\gamma_1$	0.8149 (9.96)	0.7951 (13.32)	0.7870 (14.96)
$\gamma_2$	0.5593 (2.01)	0.6237 (2.18)	0.6070 (2.11)
$\psi_{pop}$	- 7,060	- 144	- 148
$\psi_{ret}$	- 2,249	- 1,791	- 1,743
RMSE <sub>pop</sub>	93	31	43
RMSE <sub>ret</sub>	47	50	48
PRMSE <sub>pop</sub>	15.2%	5.1%	7.0%
PRMSE <sub>ret</sub>	70.1%	74.6%	71.6%

\*\*  $t_{209,0.05} = 1.96$



**Table 5.4 The average rents in Tippecanoe County townships in 1990**

District	area	average rent (\$/month)
District 1: Shelby Township	1. Otterbein 2. rest of Shelby Twp.	383 385
District 2: Wabash Township	1. West Lafayette city 2. rest of Wabash Twp.	482 389
District 3: Tippecanoe Township	1. Battle Ground 2. rest of Tippecanoe Twp.	345 367
District 4: Washington Township		360
District 5: Perry Township		412
District 6: Sheffield Township	1. Dayton 2. rest of Sheffield Twp.	349 382
District 7: Fairfield Township	1. Lafayette city 2. rest of Fairfield Twp.	372 419
District 8: Wea Township	1. Lafayette city 2. rest of Wea Twp.	418 367
District 9: Union Township		372
District 10: Wayne Township		277
District 11: Jackson Township		425
District 12: Randolph Township		415
District 13: Lauramie Township	1. Clarks Hill 2. rest of Lauramie Twp.	312 354

Source: 1990 Census Data, U.S. Department of Commerce.

**Table 5.5 Estimated coefficients for the existing housing supply submodel in the Lafayette area**

independent variables	Probability of owner occupancy ( $U'_{i,o}$ )		Probability of renter occupancy ( $U'_{i,r}$ )	
	Coefficients	t-statistics	Coefficients	t-statistics
constant	1.073790	(6.99)	- 0.147620	(-0.99)
monthly rent ( $R_i$ )	- 0.000793	(-1.91)	0.000875	(2.18)
location dummy ( $X'_{i,l}$ )	- 0.250590	(-6.31)	0.236790	(6.16)
$\bar{R}^2$	0.76783		0.76882	

\*\*  $t_{19,0.05} = 2.09$  ;  $t_{19,0.10} = 1.73$

**Table 5.6 Year 2010 population forecasts for Policy Test 1**

District	1993	APC forecast	base run	iteration 1	iteration 2	iteration 3
1	2,010	2,061	1,787	1,769	1,769	1,769
2	50,045	51,920	55,053	49,569	49,608	49,608
3	5,229	5,389	3,622	4,286	4,291	4,291
4	1,983	1,934	1,119	1,345	1,347	1,347
5	3,717	5,709	4,571	4,816	4,816	4,816
6	3,567	3,375	3,765	3,656	3,660	3,660
7	48,178	48,188	53,394	50,634	50,652	50,647
8	15,985	23,613	20,495	25,184	25,165	25,165
9	1,378	1,612	1,244	2,047	2,041	2,041
10	1,014	988	650	1,207	1,204	1,204
11	753	736	360	769	767	767
12	1,136	1,108	616	1,234	1,231	1,231
13	1,420	1,747	1,700	1,834	1,835	1,834
Total	136,415	148,380	148,376	148,376	148,386	148,382

**Table 5.7 Differences in Year 2010 population forecasts for Policy Test 1**

District	base run - APC forecast		Test 1 iteration 3 - APC forecast		Test 1 iteration 3 - base run	
	difference	percent	difference	percent	difference	percent
1	-274	-13.3 %	-292	-14.2 %	-18	-1.0 %
2	3,133	6.0 %	-2,312	-4.4 %	-5,445	-9.9 %
3	-1767	-32.8 %	-1098	-20.4 %	669	18.5 %
4	-815	-42.1 %	-587	-30.3 %	228	20.4 %
5	-1138	-19.9 %	-893	-15.6 %	245	5.4 %
6	390	11.5 %	285	8.4 %	-105	-2.8 %
7	5,206	10.8 %	2,459	5.1 %	-2,747	-5.1 %
8	-3,118	-13.2 %	1,552	6.6 %	4,670	22.8 %
9	-368	-22.7 %	429	26.6 %	797	64.1 %
10	-338	-34.2 %	216	21.8 %	554	85.2 %
11	-376	-51.1 %	29	3.9 %	407	111.1 %
12	-492	-44.4 %	123	11.1 %	615	99.9 %
13	-47	-2.7 %	87	4.9 %	134	7.9 %

**Table 5.8 Year 2010 Retail employment forecasts for Policy Test 1**

District	1993	APC forecast	base run	Test 1 iteration 1	Test 1 iteration 2	Test 1 iteration 3
1	21	21	1	0	0	0
2	3,350	3,900	5,101	4,083	4,086	4,086
3	80	380	185	131	131	131
4	29	29	6	3	3	3
5	3	3	1	0	0	0
6	71	71	11	9	9	9
7	9,703	12,628	11,990	12,912	12,910	12,910
8	371	871	679	822	821	821
9	5	5	2	3	3	3
10	12	12	3	7	7	7
11	4	4	0	0	0	0
12	20	20	6	13	13	13
13	50	50	4	4	4	4
Total	13,719	17,994	17,989	17,987	17,987	17,987

**Table 5.9 Differences in Year 2010 retail employment forecasts for Policy Test 1**

District	base run - APC forecast		Test 1 iteration 3 - APC forecast		Test 1 iteration 3 - base run	
	difference	percent	difference	percent	difference	percent
1	-20	-	-21	-	-1	-
2	1,201	30.8%	186	4.7%	-1,015	-19.9%
3	-195	-51.3%	-249	-65.5%	-54	-32.1%
4	-23	-	-26	-	-3	-
5	-2	-	-3	-	-1	-
6	-60	-	-62	-	-2	-
7	-638	-5.0%	282	2.2%	920	7.7%
8	-192	-22.0%	-50	-5.7%	142	20.9%
9	-3	-	-2	-	1	-
10	-9	-	-5	-	4	-
11	-4	-	-4	-	0	-
12	-14	-	-7	-	7	-
13	-46	-	-46	-	0	-

**Table 5.10 Comparisons of estimated average travel time based on Policy Test 1 and base case**

	Wabash Twp. (District 2) To Wea Twp. (District 8) (minutes)	Wabash Twp. (District 2) To Randolph Twp. (District 12) (minutes)	Fairfield Twp. (District 7) To Wea Twp. (District 8) (minutes)	Fairfield Twp. (District 7) To Randolph Twp. (District 12) (minutes)
1995	15	22	10	17
2010 (base case)	22	30	19	27
2010 (Test 1)	16	23	12	20

\* The travel time in 1995 was surveyed by Andrew Yen.

\* The travel time from District 7 to Districts 8 and 12 is estimated based on the traffic using old US 231.

\* The travel time from District 2 to Districts 8 and 12 is estimated based on the traffic using new US 231.

**Table 5.11 Peak hour Volume/Capacity ratio on old US 231 in 1993 and in 2010 based on Policy Test 1**

	1993		2010	
	V/C	LOS	V/C	LOS
old US 231 (between SR 26 and SR 25)	0.83	D	0.66	C
old US 231 (section in common with SR 25)	0.95	E	0.73	C
old US 231 (between SR 25 and County Rd. 500 S)	0.73	C	0.54	B
old US 231 (between County Rd. 500 S and SR 28)	0.53	B	0.56	B

**Table 5.12 Year 2010 Population forecasts for Policy Test 2**

District	1993	APC forecast	Test 2 iteration 1	Test 2 iteration 2	Difference: iteration 2 - APC forecast
1	2,010	2,061	1,765	1,765	-296
2	50,045	51,920	49,503	49,503	-2,417
3	5,229	5,389	4,276	4,276	-1,113
4	1,983	1,934	1,341	1,341	-593
5	3,717	5,709	5,059	5,059	-650
6	3,567	3,375	3,638	3,639	264
7	48,178	48,188	50,664	50,666	2,478
8	15,985	23,613	25,075	25,075	1,462
9	1,378	1,612	2,035	2,035	423
10	1,014	988	1,201	1,201	213
11	753	736	765	765	29
12	1,136	1,108	1,227	1,227	119
13	1,420	1,747	1,827	1,827	80
Total	136,415	148,380	148,376	148,379	

**Table 5.13 Year 2010 Retail employment forecasts for Policy Test 2**

District	1993	APC forecast	Test 2 iteration 1	Test 2 iteration 2	Difference: iteration 2 - APC forecast
1	21	21	0	0	-21
2	3,350	3,900	4,081	4,081	181
3	80	380	131	131	-249
4	29	29	3	3	-26
5	3	3	0	0	-3
6	71	71	9	9	-62
7	9,703	12,628	12,919	12,917	289
8	371	871	820	820	-51
9	5	5	3	3	-2
10	12	12	7	7	-5
11	4	4	0	0	-4
12	20	20	13	13	-7
13	50	50	4	4	-46
Total	13,719	17,994	17,986	17,987	

**Table 5.14 Peak hour Volume/Capacity ratio on SR 26 in 1993 and 2010 based on Policy Test 2**

	1993		2010	
	V/C	LOS	V/C	LOS
SR 26 (US 52 to Creasy Lane)	0.84	D	0.88	D
SR 26 (Creasy Lane to I-65)	0.75	C	0.90	E
SR 26 (I-65 to County Rd. 550 E)	0.59	B	0.77	D
SR 26 (County Rd. 550 E to County Rd. 700 E)	0.46	B	0.70	C
McCarty Lane (US 52 to Creasy Lane)	0.42	B	0.55	C
Haggerty Lane (US 38 to I-65)	0.45	B	0.65	C
Haggerty Lane (I-65 to County Rd. 700 E)	0.35	B	0.50	B

**Table 5.15 Year 2015 population forecasts for Policy Test 3**

District	1993	APC forecast	base run	Test 3 iteration 1	Test 3 iteration 2
1	2,010	2,061	1,766	1,873	1,873
2	50,045	52,698	49,452	49,926	49,930
3	5,229	5,389	4,261	4,159	4,167
4	1,983	1,934	1,340	1,316	1,320
5	3,717	7,122	6,013	5,920	5,932
6	3,567	3,598	4,320	4,276	4,276
7	48,178	48,532	51,294	50,595	50,595
8	15,985	23,821	25,603	25,883	25,871
9	1,378	1,813	2,183	2,227	2,223
10	1,014	988	1,190	1,214	1,211
11	753	736	776	792	791
12	1,136	1,108	1,448	1,478	1,476
13	1,420	1,747	1,901	1,882	1,881
Total	136,415	151,547	151,547	151,541	151,546

**Table 5.16 Differences in Year 2015 population forecasts for Policy Test 3**

District	base run - APC forecast		Test 3 iteration 2 - APC forecast		Test 3 iteration 2 - base run	
	difference	percent	difference	percent	difference	percent
1	-295	-14.3 %	-188	-9.1 %	107	6.1 %
2	-3,246	-6.1 %	-2,768	-5.2 %	478	0.9 %
3	-1,128	-20.9 %	-1,222	-22.7 %	-94	-2.2 %
4	-594	-30.7 %	-614	-31.7 %	-20	-1.5 %
5	-1,109	-15.6 %	-1,190	-16.7 %	-81	-1.3 %
6	722	20.1 %	678	18.8 %	-44	-1.0 %
7	2,762	5.7 %	2,063	4.2 %	-699	-1.4 %
8	1,782	7.5 %	2,050	8.6 %	268	1.0 %
9	370	20.4 %	410	22.6 %	40	1.8 %
10	202	20.4 %	475	48.1 %	21	1.7 %
11	40	5.4 %	55	7.5 %	15	1.9 %
12	340	30.7 %	368	33.2 %	28	1.9 %
13	154	8.8 %	134	7.7 %	-20	-1.1 %

**Table 5.17 Year 2015 Retail employment forecasts for Policy Test 3**

District	1993	APC forecast	base run	Test 3 iteration 1	Test 3 iteration 2
1	21	21	0	0	0
2	3,350	4,000	4,101	4,073	4,076
3	80	380	131	128	128
4	29	29	3	3	3
5	3	3	1	1	1
6	71	71	10	10	10
7	9,703	12,878	13,587	13,573	13,570
8	371	1,271	872	922	922
9	5	5	3	3	3
10	12	12	7	7	7
11	4	4	0	0	0
12	20	20	13	14	14
13	50	50	5	5	5
Total	13,719	18,744	18,733	18,739	18,739



**Table 5.18 Differences in Year 2015 retail employment forecasts for Policy Test 3**

District	base run - APC forecast		Test 3 iteration 2 - APC forecast		Test 3 iteration 2 - base run	
	difference	percent	difference	percent	difference	percent
1	-21	-	-21	-	-0	-
2	101	2.5 %	76	1.9 %	-25	-0.6 %
3	-248	-65.5 %	-249	-66.3 %	-3	-
4	-26	-	-26	-	0	-
5	-2	-	-2	-	0	-
6	-61	-	-61	-	0	-
7	709	5.5 %	282	5.4 %	-17	-0.1 %
8	-399	-31.4 %	-50	-27.4 %	50	5.7 %
9	-2	-	-2	-	0	-
10	-5	-	-5	-	0	-
11	-4	-	-4	-	0	-
12	-7	-	-6	-	1	-
13	-45	-	-45	-	0	-

**Table 5.19 Comparisons of estimated average travel time based on Policy Test 3 and base case**

	Wabash Twp. (District 2) To Wea Twp. (District 8) (minutes)	Wabash Twp. (District 2) To Randolph Twp. (District 12) (minutes)	Fairfield Twp. (District 7) To Wea Twp. (District 8) (minutes)	Fairfield Twp. (District 7) To Randolph Twp. (District 12) (minutes)
1995	15	22	10	17
2010 (Test 1)	16	23	12	20
2015 (base case)	17	24	14	22
2015 (Test 3)	15	22	13	21

\* The travel time in 1995 was surveyed by Andrew Yen.

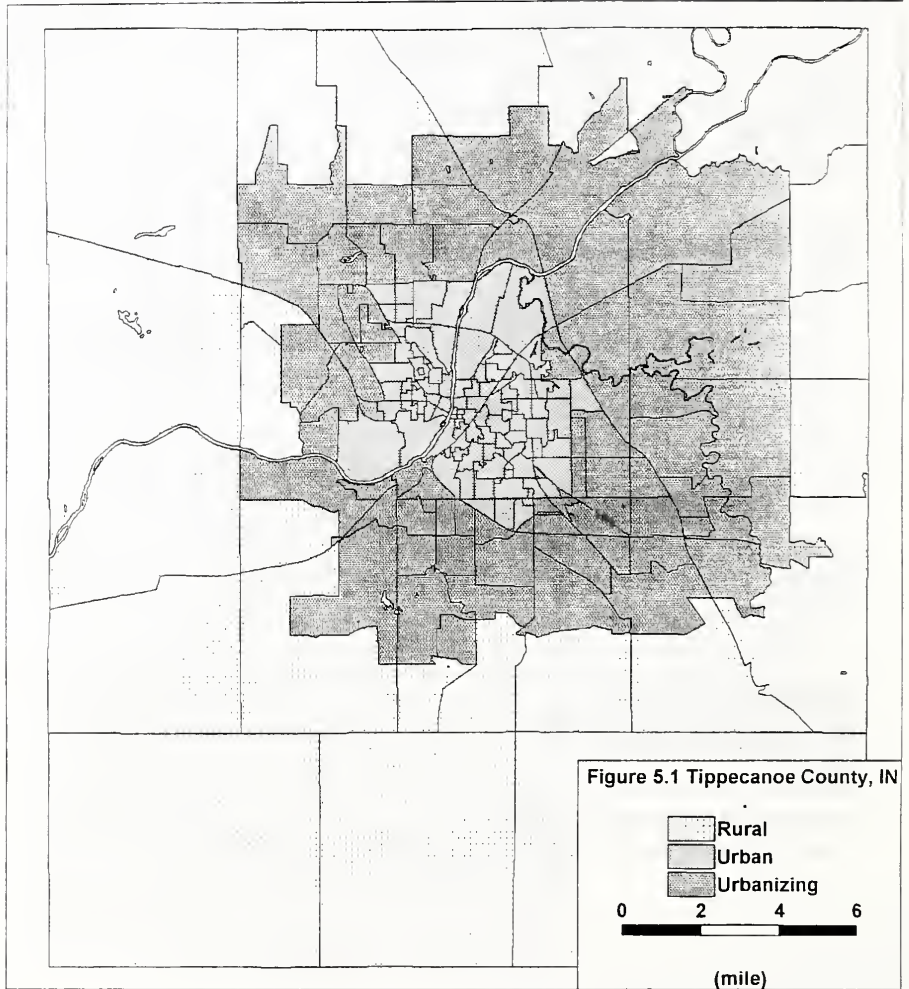
\* The travel time from District 7 to Districts 8 and 12 is estimated based on the traffic using old US 231.

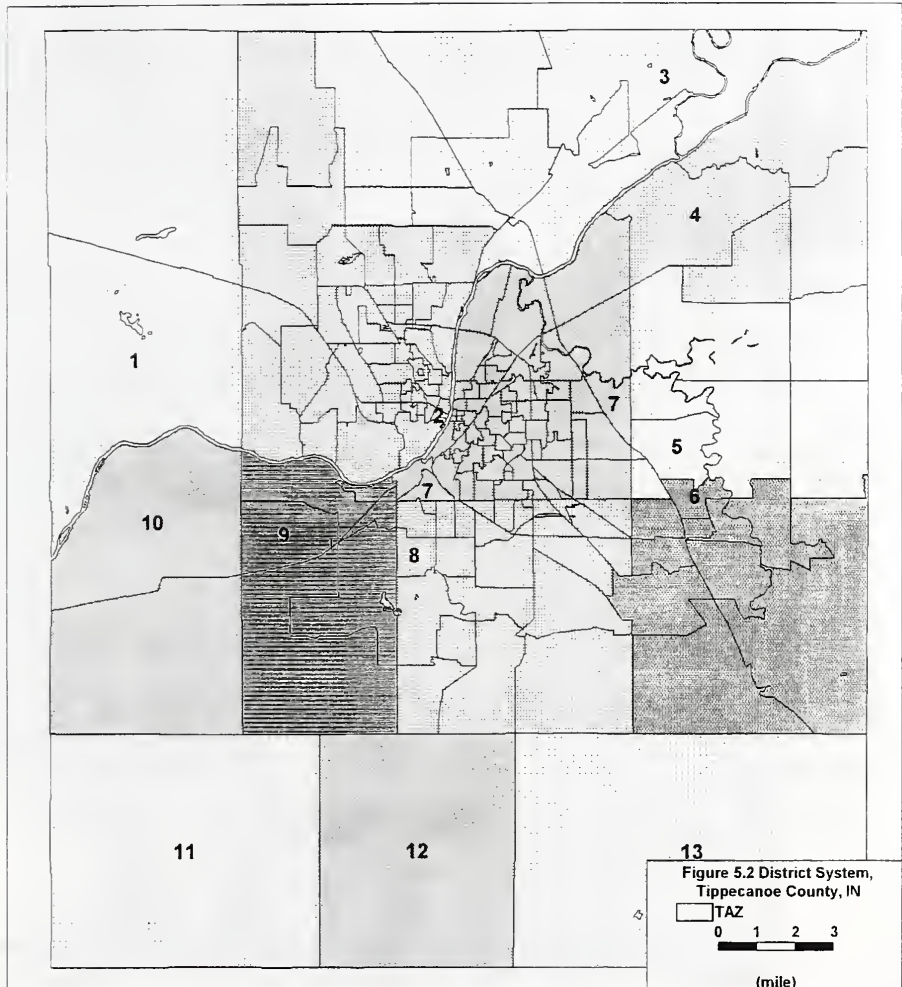
\* The travel time from District 2 to Districts 8 and 12 is estimated based on the traffic using new US 231.

**Table 5.20 Peak hour Volume/Capacity ratio on Northwestern Avenue and US 52 in 1993 and in 2015 based on Policy Test 3**

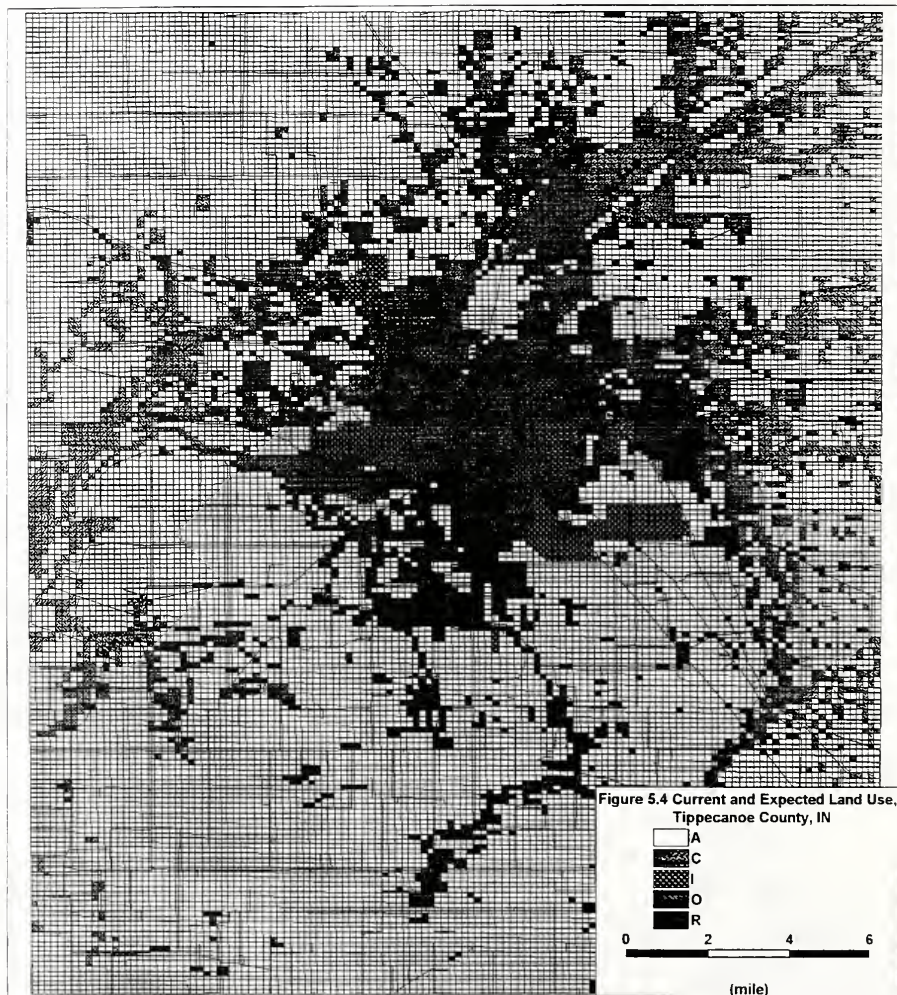
	1993		2015	
	V/C	LOS	V/C	LOS
Northwestern Avenue * (between SR 26 and US 52)	0.80	D	0.72	C
US 52 * (between Northwestern Avenue and McCormick Road)	0.70	C	0.65	C

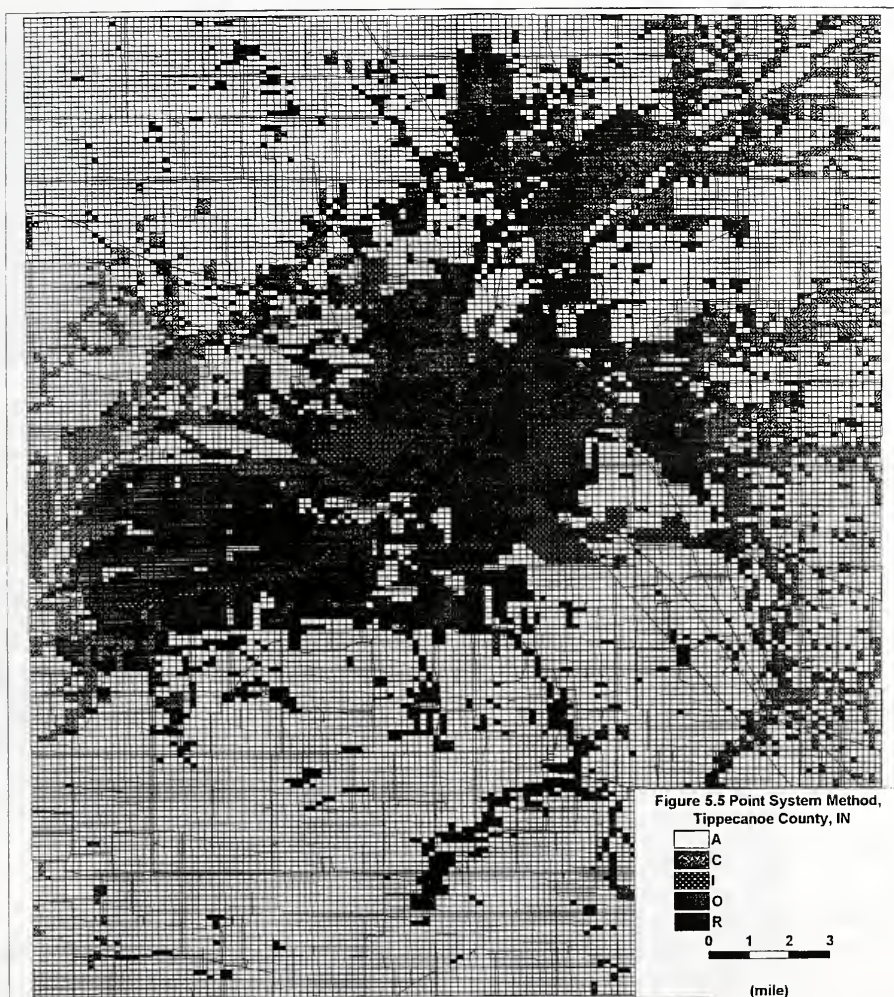
\* Northwestern Avenue and US 52 are shown in Figure 5.11.



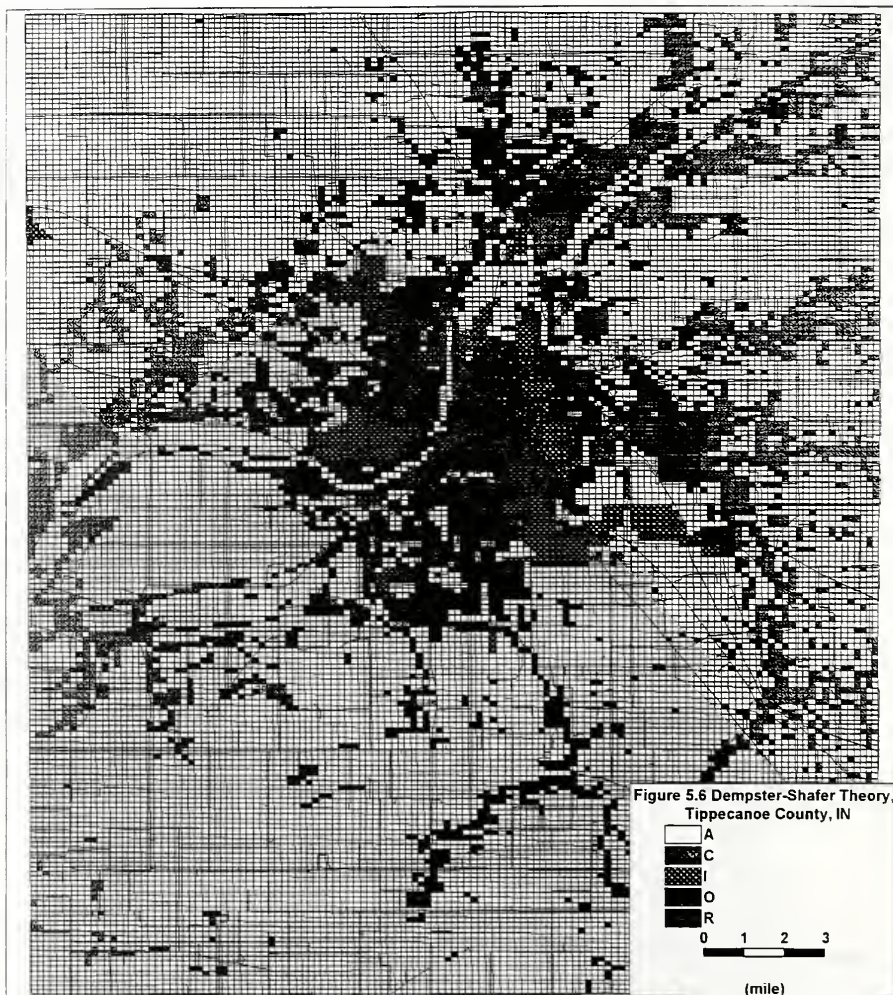




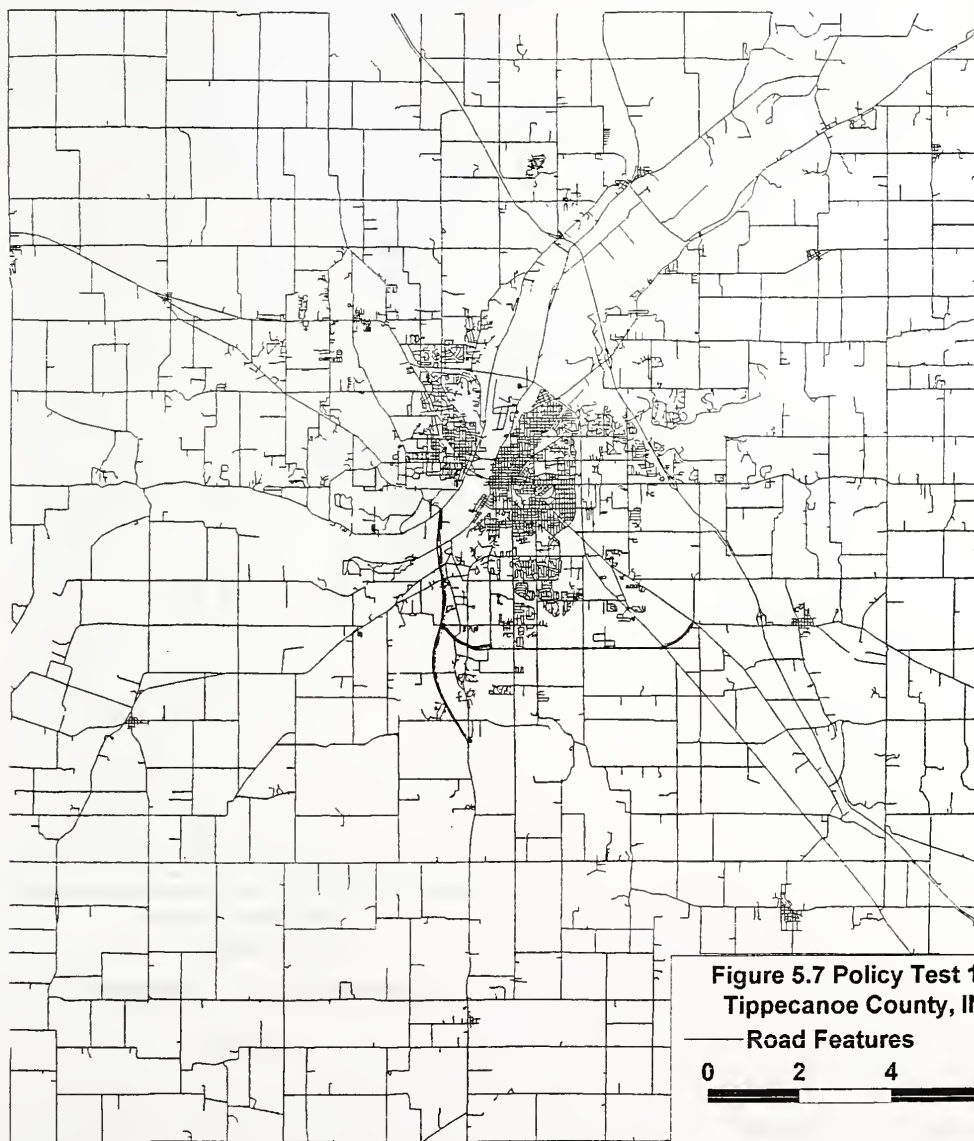


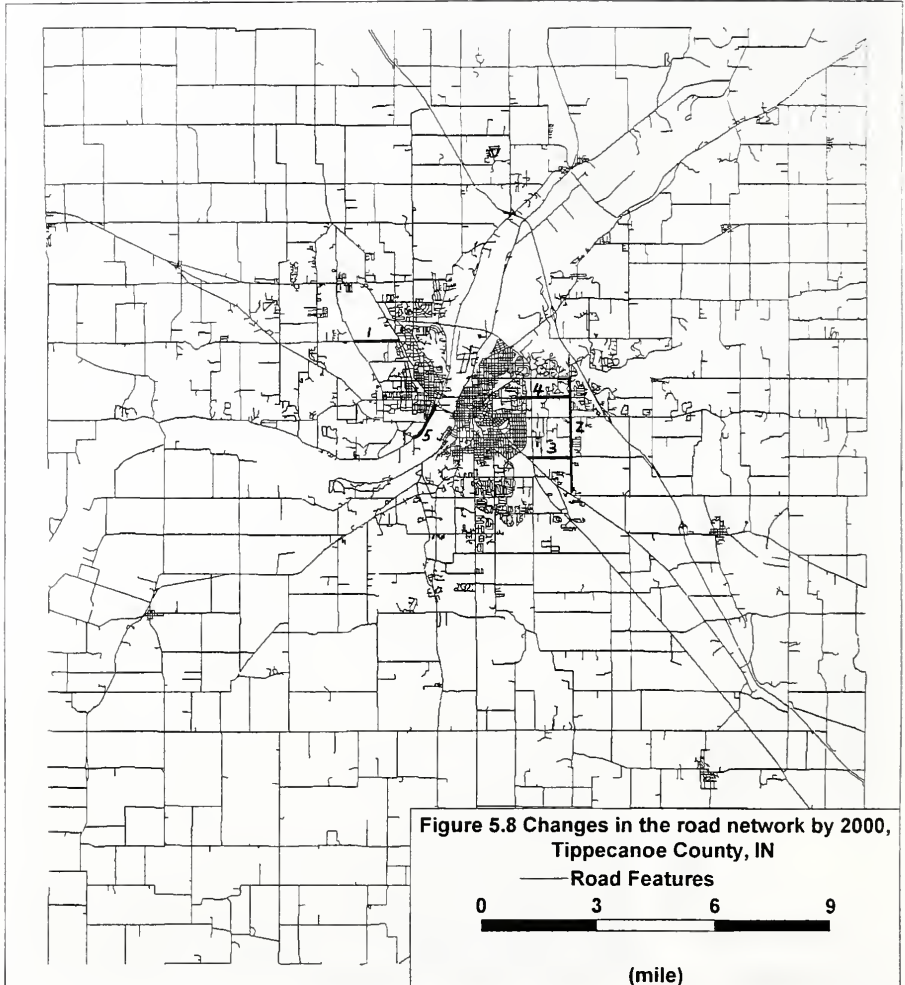


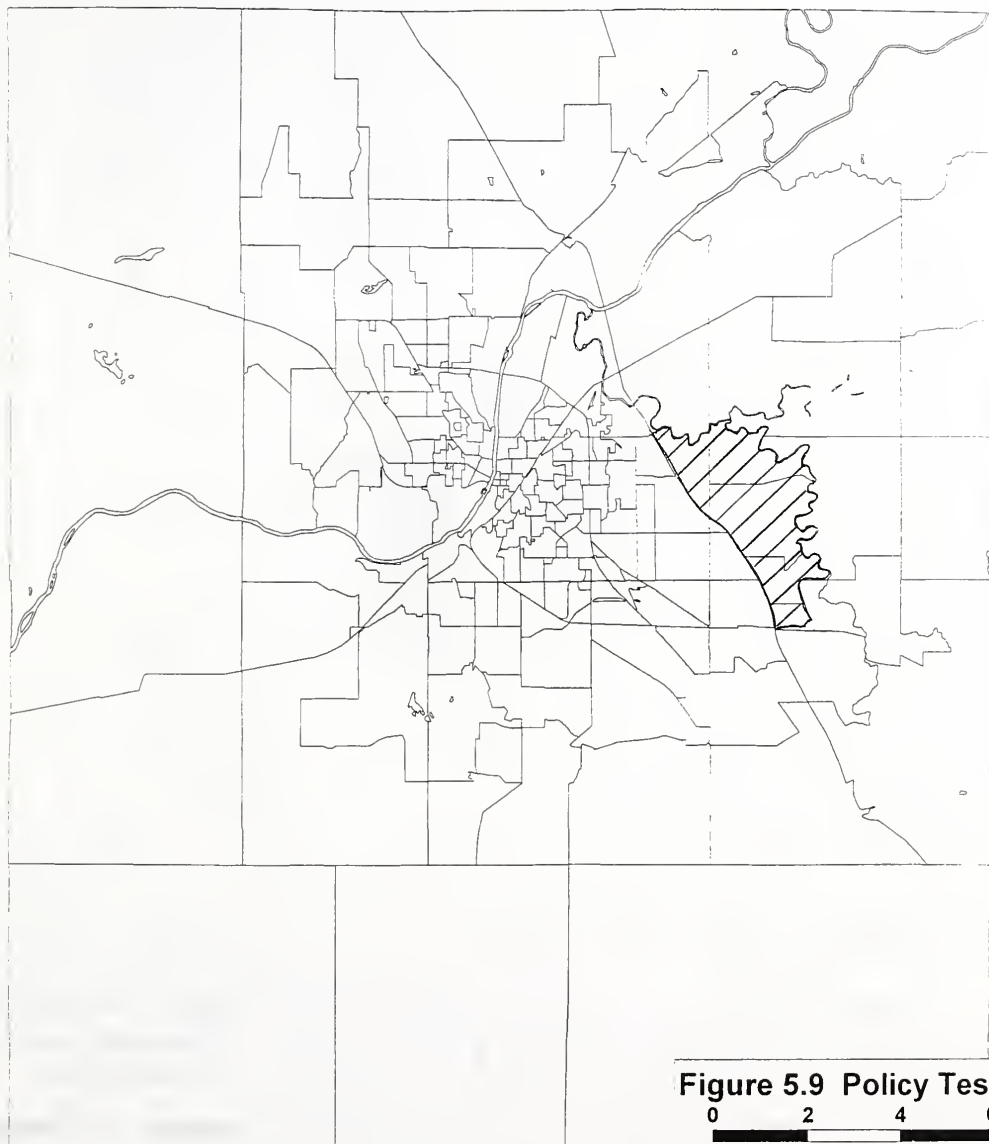


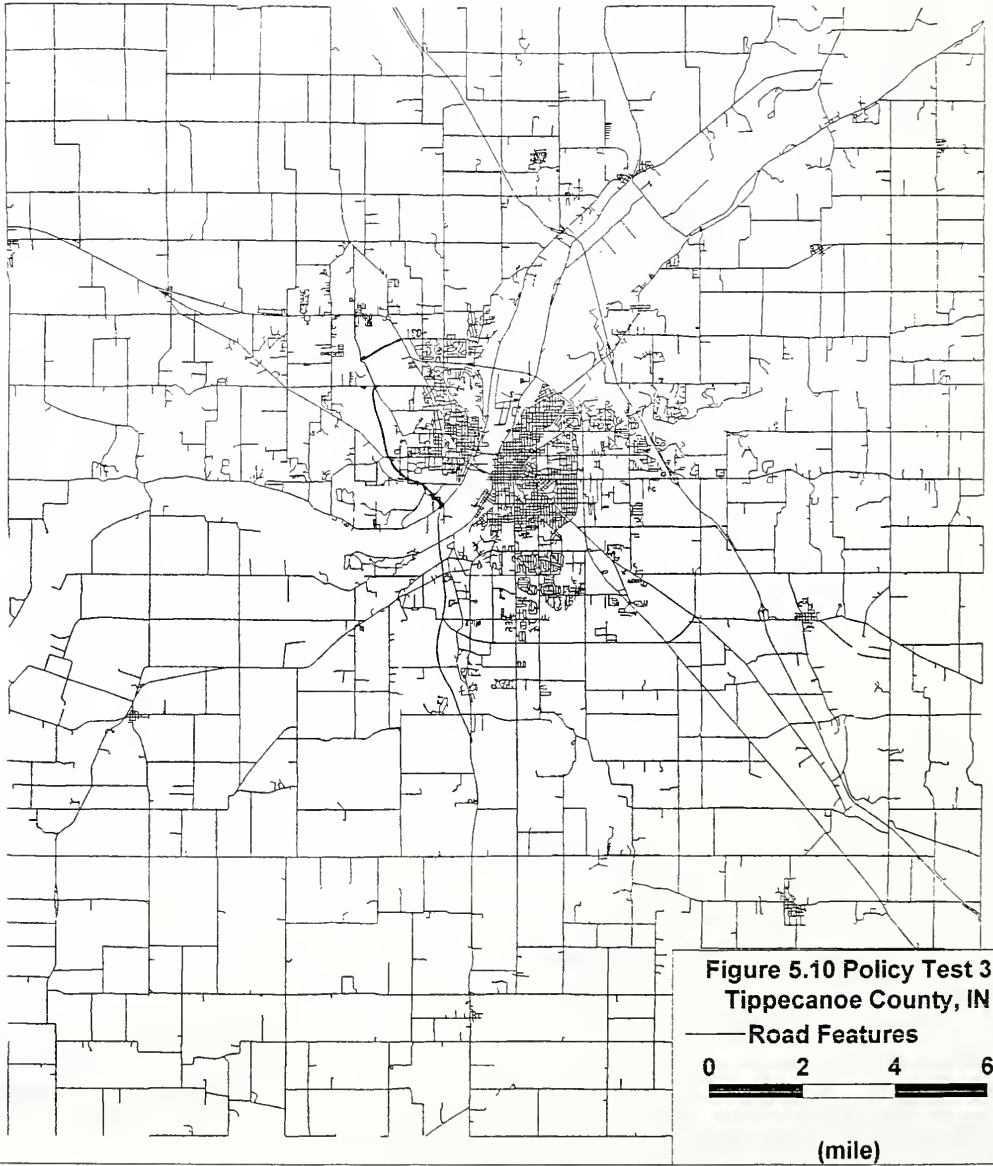


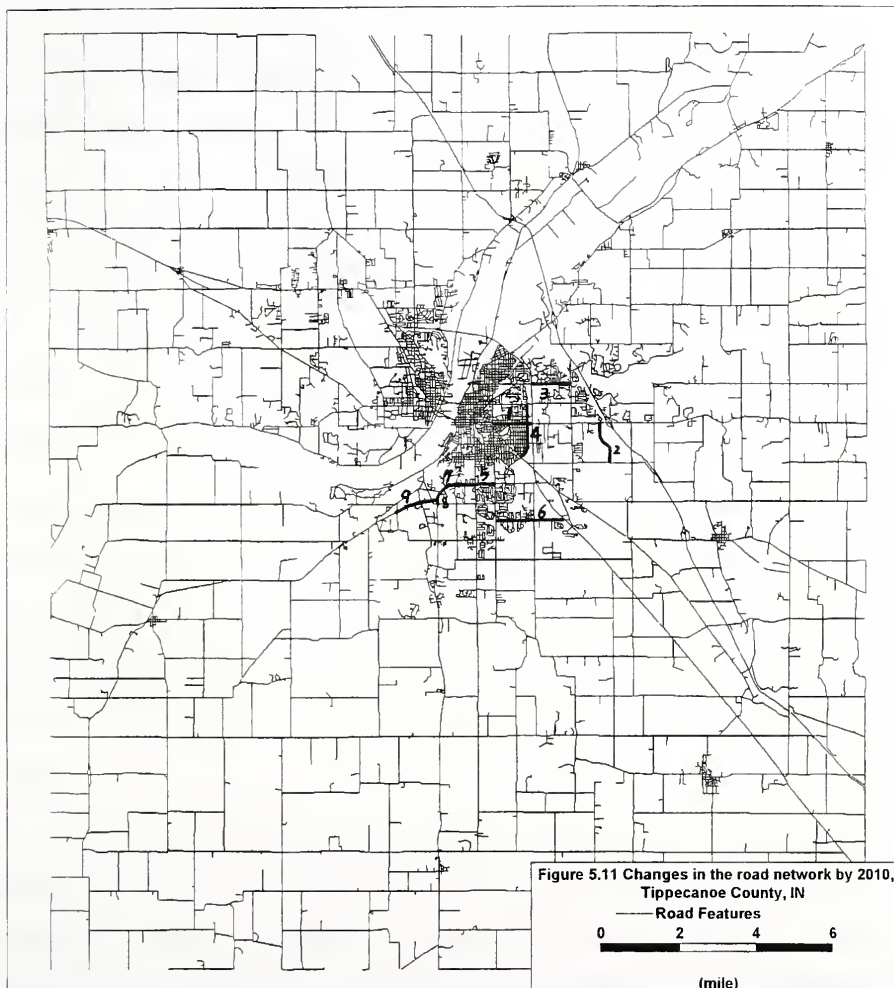














## CHAPTER 6

### CONCLUSIONS AND RECOMMENDATIONS

#### 6.1 Conclusions and Recommendations

Any operational integrated Land-Use/Transportation model is subject to various sources of error and bias. These could exist in the data, in the mathematical form, and in the assumptions used in the analysis. Such sources of error and bias are unavoidable (Anas and Duann, 1985). Therefore, in order to keep the error or bias as small as possible, an appropriate integrated Land-Use/Transportation model should be based on good modeling practice, realistic assumptions, legitimate calibration, and a recognition of the data available to the user. Furthermore, an integrated model should be able to help planners quantify impacts caused by long range transportation plans on land use. At the same time, the model can help planners to mitigate congestion and environmental impacts that result from the interaction between land use and transportation.

The land use allocation module in this research consists of a residential location model, an employment location model, a land consumption model, and a land use potential model. The residential location and employment location models are analogous to DRAM/EMPAL proposed by Putman (1983,1996). The two models can be used to derive the spatial distributions of population and retail (or nonbasic) employment over the study area, based on random utility theory at the household level, if such data are available.

In order to overcome DRAM/EMPAL's (Putman, 1983) lack of the land market mechanism, Putman (1995) incorporated a procedure (LANCON) to calculate land consumption. Because this procedure proposed by Putman makes only a crude representation of the balance between the demand and supply of residential housing choices, a significantly different land consumption approach based on micro-economic theory is incorporated into the integrated model in this research. This land consumption model can simulate the profit-maximizing behavior of housing or land owners over time.



The major purpose of the land consumption model is to satisfy the need to reach a balance between demand and supply in the housing market during each time period.

The land consumption model used in this research is a modified version of Anas's work in CATLAS (Anas and Duann, 1985). However, the work of CATLAS does not deal with employment location. In general, households and employers compete for location, and employers often outbid households for location where they compete (Waddell, 1995). Hence, the validity of the predictions of land price impacts on changes in land use or transportation strategies in CATLAS will be questionable. The land use allocation module in this research simultaneously considers housing location and employment location. In order to reflect the situation of employers (or businesses) often outbidding households, the developable area in each zone is always first assigned to employment locations where households compete for the same land.

All the required data that are described in Chapter 4 can be collected or extracted from census data, or assembled by planners in each MPO. Furthermore, the integrated model developed in this research has demonstrated other properties. They are:

- 1) For the integrated model, it is assumed that the locations and amount of nonretail (or basic) employment for the forecast year in the study area are determined outside the integrated model. Population and retail employment are jointly allocated in the study area to the extent that relationships between population and retail employment are captured by available data.
- 2) The work of building a GIS database for a land use potential study is time-consuming. However, a GIS database, once built, will be easily modified to satisfy the future needs of land use and transportation planning.
- 3) The integrated model allows comparison between two runs of the integrated model, one of which is a do-nothing case (or base case) and the other is a policy to be tested. The model can be used to evaluate land use policies such as new residential or commercial construction and transportation policies such as the improvement of the highway or transit network.

- 4) The travel demand module contains trip generation and trip distribution stages for estimating home-based school, home-based other, non-home-based, external-internal and external-external trips. It splits the trip matrices into individual modes using diversion curve approach (or logit model) and then assigns the vehicle trips on the links of network being analyzed using equilibrium assignment approach (or multipath stochastic assignment procedure). It includes capacity restraint on these links to represent congestion on the routes and considers the behavior of individual stochastic route choice. It has been demonstrated that a network equilibrium can be reached in the travel demand module.

From the results of an application in the Lafayette area, the integrated model developed in this study has shown the characteristics described below.

The integrated model can find a combined (and stabilized) solution for the land-use and transportation modules. It can reflect the effects of changes in the travel network on the level of activity -- for example, more population and retail employment will be attracted to locations that have greater accessibility to workplaces, shopping, etc. The model can also show the impacts of changes in land use patterns on travel patterns -- for example, more traffic will be generated because the current agricultural use of a region is converted into residential, commercial, or industrial use.

Furthermore, it has been demonstrated that the integrated model can be used to quantify positive or negative effects of long range transportation plans (e.g., the improvement or new construction of highway network) on land use (e.g., the spatial distribution of population and employment). This will satisfy the requirement of the ISTEA of 1991. (See Section 5.1.5.)

In order to satisfy the requirement of the CAAA of 1990, any operational integrated model should be able to help planners mitigate environmental impacts (e.g., traffic-induced air pollution) generated from the interaction between land use and transportation. Actually, the output from the travel demand model in this study (speed, link volume, V/C, travel time) can be easily input to an emissions model. Thus, for example, MPOs in non-

attainment areas can use the results generated from the emissions model to indicate whether new transport or land use projects will have positive or negative impacts on air quality. Furthermore, the information (e.g., the changes in traffic conditions) provided by the integrated model can help decision makers or planners devise strategies to avoid negative future effects.

One major benefit of the land use potential study is that the information generated can serve as a data base for a variety of planning activities, including the eventual land use plan itself. However, it must be stressed that the data generated is not the plan, but rather its foundation.

Because the land use potential study model is built based on a GIS database, this model can manipulate large amount of data more easily than an individual can. Furthermore, the point system approach and the Dempster-Shafer evidence theory can be applied over a large study area, and the land use potentials generated can be counted on to sort out competing land uses rationally and consistently. The procedures proposed in the land use potential study have a technical basis and are systematically quantitative, rather than intuitive. Hence, such a study can be readily replicated or updated in the future.

The findings from the present study demonstrate that the land use potential study can produce accurate breakdowns of current land area and rational estimates of future land area for residential, commercial, industrial, agricultural and open space uses. The areas for residential, commercial and industrial development are used as the input to the land use allocation module. Test runs for the Lafayette area has demonstrated that accurate calculations of residential and commercial area can improve the accuracy of parameter calibration in the residential location model and in the employment location model, respectively. (See Section 5.1.1)

Some current limitations of the integrated model developed in this study are:

- 1) Because of data constraints, the data on households (or population) and dwelling units are collected for aggregate groups (e.g., high, middle, and low income) and the data

on employment are divided into several categories (for example, retail, manufacturing and other employment). Therefore, the integrated model does not capture the individual's choice process with respect to housing and workplace location if it is based on aggregate data.

- 2) Although the integrated Land-use/Transportation model produces believable results for the transportation impacts of land use changes, there are some indications that the model overestimates the land use impacts of transportation changes. Of course, land use changes are made in response to a variety of stimuli, only one of which is transportation. In focusing on the transportation stimulus, the calibration and use of the integrated model may become too sensitive to transport network changes. Users of the model should keep this in mind, and place more credence on the land use trends rather than the specific forecasted values.
- 3) The movement of goods is closely related to the level of economic activities – for example, retail and nonretail. In the current study, the nonretail employment is exogenously determined from other resources or studies. Therefore, only the movement of passengers is considered in the integrated model. The integrated model can incorporate truck trips to represent the movement of goods, if such data are available.

## 6.2 Future Research

The integrated model developed in this study only determines the amount and location of household and retail (or nonbasic) employment. It is assumed that nonretail (or basic) employment is determined outside the integrated model. Actually, household, retail (or nonbasic) and nonretail (or basic) employment may be jointly determined. Therefore, further research based on the current study might benefit from attempting to capture the interactions between residence, retail and nonretail activities (including the movement of goods). (Refer to the equations (3.4) and (3.13) in Section 3.2.)

Because available data are in aggregate form, the integrated model developed in this research only applies an aggregate process to capture the interaction between land use and

transportation. An aggregate model does not realistically reflect the role of household, businesses, developers, or governments in making decisions concerning the development process, for example, location choices and investment. Furthermore, an aggregate land-use/transportation models cannot explicitly explain the links between the decision processes (housing choice, mode choice, route choice and so forth) of individuals over time (Waddell, 1995).

Actually, the integrated model developed in this research can be easily modified to simulate the individual's process for housing and workplace choices if the data are available in disaggregate form. For this reason, the application of disaggregate microsimulation approaches to land-use/transportation models would be encouraged in any future study.

### **6.3 Implementation**

As the project ended, the integrated model's software was demonstrated to interested MPOs and INDOT personnel. The project's implementation report will enable users to prepare data files in the required formats and carry out each step properly.

## LIST OF REFERENCES

- Alonso, William. 1964. Location and Land Use, Cambridge, MA: Harvard University Press.
- Armstrong, B. D. 1977. The Need for Route Guidance, Department of the Environment, Department of Transport, Transport and Road Research Laboratory Supplementary Report SR330.
- Anas, Alex. 1981. "The Estimation of Multinomial Logit Models of Joint Location and Travel Mode Choice from Aggregate Data", Journal of Regional Science 21, No.2, 223-242.
- Anas, Alex. 1982. Residential Location Markets and Urban Transportation, New York: Academic Press.
- Anas, Alex. 1983. "Discrete Choice Theory, Information Theory and the Multinomial Logit and Gravity Models", Transportation Research 17B, 13-23.
- Anas, Alex and Liang Shyong Duann. 1985. "Dynamic Forecasting of Travel Demand, Residential Location and Land Development", Paper of the Regional Science Association 56, 37-58.
- APC (Area Plan Commission) 1981. The Comprehensive Plan for Tippecanoe County, Volume 1: The Land Use Potentials Study Technical manual, Area Plan Commission, Tippecanoe County, Indiana.
- APC (Area Plan Commission) 1992. Unified Zoning Ordinance, Area Plan Commission, Tippecanoe County, Indiana.
- Aronoff, Stan. 1989. Geographic Information Systems: A Management Perspective, WDL Publications, Ottawa, Ontario.
- Batty, M. 1992. "Urban Modeling in Computer-Graphic and Geographic Information System Environment", Environment and Planning B 19, 663-688.
- Baxter, R. and I. Williams. 1975. "An Automatically Calibrated Urban Model", Environment and Planning A 17, 3-20.
- Bazaraa, Mokhtar S., Hanif D. Sherali and C. M. Shetty. 1993. Nonlinear Programming: Theory and Algorithms, John Wiley & Sons, Inc.
- Ben-Akiva, M. and S. R. Lerman. 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, MIT Press, Cambridge, MA.



- Beckmann, M., C. McGuire and C. Winston. 1956. Studies in the Economics of Transportation, Yale University Press, New Haven.
- Benshoof, J. A. 1970. "Characteristics of Drivers' Route Selection Behavior", Traffic Engineering and Control 11, 604-606.
- Berechman, J. and K. A. Small. 1988. "Research Policy and Review 25. Modeling Land Use and Transportation: An Interpretive Review for Growth Areas", Environment and Planning A 20, 1285-1309.
- Chang, Y. C. 1991. "A Dempster-Shafer Reasoning Shell for Spatial Indexing", Master Thesis, Purdue University.
- Cohen, M. S. 1986. "An Expert System Framework for Non-monotonic Reasoning about Probabilistic Assumptions", Uncertainty in Artificial Intelligence, L. N. Kanal and J. F. Lemmer eds., New York, Elsevier Science Publishers, B.V.
- Cochrane, R. A. 1975. "A Possible Economic Basis for the Gravity Model", Journal of Transport Economics and Policy 9, 34-49.
- Dafermos, S. C. 1972. "The Traffic Assignment Problem for Multiclass-User Transportation Network", Transportation Science 6, 73-87.
- de la Barra, Tomas. 1989. Integrated Land Use and Transport Modeling, Cambridge University Press, Cambridge.
- Dempster, A. P. 1967. "Upper and Lower Probabilities Induced by a Multivalued Mapping", Ann. Math. Statist. 38, 325-339.
- Dempster, A. P. and A. Kong. 1988. "Uncertain Evidence and Artificial Analysis", Journal of Statistical Planning and Inference 20, 355-368.
- Dial, R. B. 1971. "A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration", Transportation Research 5, 83-112.
- Dickey, J. W. 1983. Metropolitan Transportation Planning, McGraw-Hill, New York.
- Domencich, T. A. and Daniel McFadden. 1975. Urban Travel Demand: A Behavioral Analysis, North-Holland, Amsterdam.
- Echenique, M. H., A. D. J. Flowerdew, J. D. Hunt., T. R. Mayo, I. J. Skidmore and D. C. Simmonds. 1990. "The MEPLAN Models of Bilbao, Leeds and Dortmund", Transport Reviews 10, 309-322.



Evans, S. P. 1976. "Derivation and Analysis of Some Models for Combining Trip Distribution and Assignment", Transportation Research 10, 37-57.

Fisk, C. S. 1980. "Some Developments in Equilibrium Traffic Assignment", Transportation Research B 14B, 243-255.

Garin, R. A. 1966. "A Matrix Formulation of the Lowry Model for Intrametropolitan Activity Location", Journal of American Institute of Planners 32, 361-364.

Giarratano, Joseph and Gary Riley. 1994. Expert Systems: Principles and Programming. PWS Publishing Company, Boston.

Giuliano, Genevieve. 1995. "The Weakening Transportation-Land Use Connection", Access No. 6, 3-11, University of California Transportation Center.

Goldstein, G. S. and T. J. Gronberg. 1984. "Economics of Scope and Economies of Agglomeration", Journal of Urban Economics 16, 91-104.

Gordon, J. and E. H. Shortliffe. 1985. "A Method for Managing Evidential Reasoning in a Hierarchical Hypothesis Space", Artificial Intelligence 26, 323-357.

Hamerslag, Rudi, Eric C. Van Berkum, and Michael A. Replogle. 1993. "A Model to Predict the Influence of New Railways and Roads on Land Use Development", presented at the 72nd Annual Meeting of the Transportation Research Board, Washington, DC.

Hamerslag, Rudi, Eric C. Van Berkum, and Michael A. Replogle. 1994. "A Dynamic Model with Elastic Constraints to Predict the Influence of Railways, Roads and Congestion on Land Use Development in Metropolitan Areas", accepted for publication in Transportation Research Record 1994.

Harvey, G. and E. Deakin. 1993. "A Manual of Regional Transportation Modeling Practice for Air Quality Analysis", National Association of Regional Councils, Version 1.0.

Hawthorn, Gary. 1991. "Transportation Provisions in the Clean Air Act Amendments of 1990", ITE Journal 20, 17-24.

Henderson James M. and Richard E. Quandt. 1980. Microeconomic Theory: A Mathematical Approach, McGraw-Hill, Inc.

Horowitz, Alan J. 1993. "Reference Manual Highway Land-Use Forecasting Model II+ for Windows, Version 2.0", Draft, University of Wisconsin -- Milwaukee.

Hunt, J. D. and D. C. Simmonds. 1993. "Theory and Application of an Integrated Land-Use and Transport Modeling Framework", Environment and Planning B: Planning and Design 20, 221-244.

Hutchinson, R. D., R. W. McNees and C. L. Dudek. 1977. "Survey of Motorist Route Selection Criteria", Transportation Research Record 643, Transportation Research Board, Washington DC, 45-48.

Ingram, G., John F. Kain and J. R. Ginn. 1972. The Detroit Prototype of the NBER Urban Simulation Model, New York: National Bureau of Economic Research.

Kain, John F. and William C. Apgar, Jr. 1985. Housing and Neighborhood Dynamics: A Simulation Study, Cambridge: Harvard University Press.

Kim, T. J. 1983. "A Combined Land Use-Transportation Model when Zonal Travel Demand is Endogenously Determined", Transportation Research B 17B, No.6, 449-462.

Kim, T. J. 1990. Advanced Transport and Spatial System Models: Application to Korea, Springer-Verlag New York Inc.

Landis, John D. 1993. "CUF Model Simulation Results: Alternative Futures for the Greater Bay Area Region", Working Paper 592, Berkeley, CA: Institute of Urban and Regional Development, University of California at Berkeley. As cited in Wegener 1994.

Laurini, Robert and Derek Thompson. 1994. Fundamentals of Spatial Information Systems, Academic Press Limited, San Diego.

Lawrence, J. D. and T. D. Garvey. 1982. "Evidential Reasoning: A Developing Concept", Proc. Int. Conf. on Cybernetics and Society, 6-9. As cited in Giarratano and Riley 1994.

Lerman, S. R. 1977. "Location, Housing and Automobile Ownership and Mode to Work: A Joint Choice Model", Transportation Research Record 610, 71-84.

Lowry, I. S. 1964. Model of Metropolis, Rand Corporation, Santa Monica, California.

Lunn, S. E. 1978. Route Choice by Drivers, Department of the Environment, Department of Transport, TRRL Report SR374, Transport and Road Research Laboratory, Crowthorne.

Mackett, Roger L. 1983. "The Leeds Integrated Land-Use Transport Model (LILT)", Supplementary Report SR 805. Crowthorne: Transport and Road Research Laboratory. As cited in Wegener 1994.

Mackett, Roger L. 1990. "Comparative Analysis of Modeling Land-Use Transport Interaction at the Micro and Macro Levels", Environment and Planning A 22, 459-475.

Mackett, Roger L. 1993. "Structure of Linkages between Transport and Land-Use", Transportation Research B 27B, No.3, 189-206.

Martinez, Francisco. 1992a. "Towards the 5-Stage Land-Use Transport Model. Land Use, Development and Globalisation", Selected Proceedings of the Sixth World Conference on Transport Research, Lyon. St.-Just-la-Pendue: Press de l'Imprimerie Chirat, 79-90. As cited in Wegener 1994.

Martinez, Francisco. 1992b. "The Bid-Choice Land-Use Model: An Integrated Economic Framework", Environment and Planning A 24, 871-885.

McFadden, D. 1973. "Conditional Logit Analysis and Qualitative Choice Behavior", In Frontiers in Econometrics, 105-140, New York: Academic Press.

McFadden, D. 1978. "Modeling the Choice of Residential location", In Spatial Interaction Theory and Planning Models, A. Karlqvist et al. eds., 105-140, Amsterdam: North-Holland Publishing.

NARC (National Association of Regional Councils) 1992. "Suit Surfaces Key Questions on the Impact of Conformity", Clean Air/Transportation Report, Washington DC.

Neuburger, H. 1971. "User Benefit in the Evaluation of Transport and Land Use Plans", Journal of Transport Economics and Policy 5, 52-75.

Orski, C. Kenneth. 1989. "A Realistic Appraisal of Traffic", Urban Land, 34-36.

Pindyck, Robert S. and Daniel L. Rubinfeld. 1991. Econometric Models and Economic Forecasts, McGraw-Hill, Inc.

Prastacos, P. 1986. "An Integrated Land-Use/Transportation Model for the San Francisco Region", Environment and Planning A 18, 307-322 and 511-528.

Putman, S. H. 1974. "Preliminary Results from an Integrated Transportation and Land Use Models Package", Transportation 3, 193-224.

Putman, S. H. 1983. Integrated Urban Models, London.

Putman, S. H. 1991. Integrated Urban Models 2: New Research and Applications of Optimization and Dynamics, Pion, London.

Putman, S. H. 1995. "EMPAL and DRAM Location and Land Use Models: An Overview", Land Use Modeling Conference, Dallas, TX.

Putman, S. H. 1996. "Extending the EMPAL and DRAM Models: The Theory - Practice Nexus", presented at the 75th Annual Meeting of the Transportation Research Board, Washington, DC.

Ratcliffe, E. P. 1972. "A Comparison of Drivers Route Choice Criteria and Those Used in Current Assignment Processes", Traffic Engineering and Control 13, 526-529.

Rho, J. H. and T. J. Kim. 1989. "Solving a Three-Dimensional Urban Activity Model of Land Use Intensity and Transport Congestion", Journal of Regional Science 29, 595-613.

Shafer, G. 1976. A Mathematical Theory of Evidence, Princeton University Press, Princeton.

Sheffi, Yosef. 1985. Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods, Prentice Hall, Englewood Cliffs, New Jersey.

Smith, M. J. 1979. "The Existence, Uniqueness and Stability of Traffic Equilibria", Transportation Research B 13B, 295-304.

Soil Survey: Tippecanoe County Indiana, 1959, United States Department of Agriculture, Soil Conservation Service, in cooperation with Purdue University Agricultural Experiment Station, U.S. Government Printing Office, Washington, DC.

Star, Jeffrey and John Estes. 1990. Geographic Information Systems: An Introduction, Prentice-Hall, Inc., New Jersey.

TETUP (Travel Estimation Techniques for Urban Planning) 1994. Travel Estimation Techniques for Urban Planning, final report prepared for National Cooperative Highway Research Program, Transportation research Board, National Research Council, Washington, DC.

USDA-SCS, 1975. Guide to Soil Survey Interpretation Sheets, USDA-SCS-Indiana.

Van Vliet, D. 1976. "Road Assignment - I. Principles and Parameters of Model Formulation", Transportation Research 10, 137-143.

Van Vliet, D. and P. C. Dow. 1979. "Capacity Restrained Road Assignment", Traffic Engineering and Control 20, 296-305.

Wachs, M. 1989. keynote address at the 41st California Transportation and Public Works Conference, Garden Grove, CA, 17-19 May 1989, as reported by Robin Nielson-Cerquone in "Looking Beyond Slow Growth", ITS Review, 4-6.

Waddell, Paul. 1995. "Integrating Land Use and Transportation Models", Working paper, draft, Land Use Model Conference at Dallas.

Wegener, Michael. 1986. "Transport Network Equilibrium and Regional Deconcentration", Environment and Planning A 18, 437-456.

Wegener, Michael, Roger L. Mackett and David C. Simmonds. 1991. "One City, Three Models: Comparison of Land-use/Transport Policy Simulation Models for Dortmund", Transport Reviews 11, 107-129.

Wegener, Michael. 1994. "Operational Urban Models: State of the Art", Journal of the American Planning Association 60, No. 1, 17-29.

Wesley, Leonard P. 1986. "Evidential Knowledge-based Computer Vision", Optical Engineering 25, 363-379.

Williams, H. C. W. L. 1976. "Travel Demand Models, Duality Relations and User Benefits Analysis", Journal of Regional Science 16, 147-166.

Wilson, Alan G. 1967. "A Statistical Theory of Spatial Distribution Models", Transportation Research 1, 253-269.

Wilson, Alan G. 1970. Entropy in Urban and Regional Modeling, London: Pion.

Zadeh, Lotfi A. 1984. "Review of Books: A Mathematical Theory of Evidence", The Artificial Intelligence Magazine, 81-83.



## **Appendix A. Land-Use Potential Study – Point System Approach**

In 1981, the Area Plan Commission (APC) of Tippecanoe County, Indiana, proposed a land-use potential decision-making model. It is conducive to open discussion and permits input from a wide range of participants. Because planning decisions are based on an explicit set of rules, not solely the judgment of one or more planners, the potential for controversy in the subsequent planning process can be reduced.

The model built by the APC is in the form of a matrix that indicates the interrelationship between various factors and land use categories (Table A.1) (APC, 1981). The model is to determine the potential for five land use categories: residential, agricultural, industrial, commercial and open space. The factors that the APC uses to determine appropriate land are explained below (APC, 1981).

- 1) Soil productivity - the relative ability of a given soil type to yield crops.
- 2) Soil limitation - the relative ability of a given soil type to withstand various kinds of development: its constructability.
- 3) Tendency to flood - whether or not a given soil type is regularly subject to ponding or flooding by stream or river.
- 4) Forestation - the presence or absence of significant numbers of trees.
- 5) Sanitary sewer availability - relative access to a trunk line known to have excess capacity.
- 6) Accessibility - a measure of proximity to major and minor roadways and their intersections.
- 7) Railroads and airports proximity - a distance measure to these major transportation facilities and their areas of influence, and
- 8) Current and expected use - a generalized indication of how land is being used at the present time, including all major development projects currently in the drawing-board stage or some form of which has been granted official approval.



These factors describe the physical characteristics of soil, land and regional attributes within the manmade infrastructure. The inclusion of a “current and expected use” factor ensures against generating any result that is contrary to reality.

In Table A.1, a range of values, from +2 to -1, is assigned to each specific interaction between factor and potential use category. The range of values corresponds to the range of interaction as follows:

+2 = highly desirable,

+1 = more than acceptable,

0 = neutral or not applicable or just acceptable, and

-1 = usually not acceptable

The APC had no particular evidence indicating that any factor was of greater influence than any others in determining potential for use within categories. Hence in this decision-making model each factor is given equal weight. However a user could choose a different set of “interaction values” or give unequal weights to the factors.

Table A.1 Land use potential decision-making matrix

Factors	Residential				Agricultural			
	2	1	0	-1	2	1	0	-1
Soil Productivity	low	m	h	vh	vh	h	m	low
Soil Limitation	slight		m	severe			all	
Tendency To Flood			all soils				all soils	
Forested		yes					yes	
Sanitary Sewer (ft)	<1000							< 1000
Accessibility	paved <.5 mi	paved >.5 mi		major ins. or R.O.W			all	
Railroads & Airports (ft)			> 300	< 300			all	
Current & Expected Use	R			O	A			O

Table A-1. (continued)

Factors	Industrial				Commercial			
	2	1	0	-1	2	1	0	-1
Soil Productivity	low	m	h	vh	low	m	h	vh
Soil Limitation	slight		m	severe	slight		m	severe
Tendency To Flood			all soils				all soils	
Forested				yes			yes	
Sanitary Sewer (ft)	<1000	1000-7920		>7920	<1000	1000-2640		>2640
Accessibility	major ins. R. O.W	paved <.5 mi	int. R.O.W	paved >.5 mi	major ins. R. O.W	paved <.5mi	int. R.O.W	paved <.5 mi
Railroads & Airports (ft)	< 300	300-2640	> 2640				> 300	< 300
Current & Expected Use	I			O	C			O.

Table A-1. (continued)

Factors	Open Space			
	2	1	0	-1
Soil Productivity			all	
Soil Limitation			all	
Tendency To Flood	flooding soils		all other soils	
Forested	yes			
Sanitary Sewer (ft)			all	
Accessibility			all	
Railroads & Airports (ft)		< 300	> 300	
Current & Expected Use	O			

m: moderate

h: high

vh: very high

R: residential

A: agricultural

I: industrial

C: commercial

O: open space



## Appendix B. Dempster-Shafer Evidence Theory

Dempster (1967) proposed a new concept of lower and upper limits on probability. He attempted to model uncertainty by a range of probabilities rather a single probabilistic number. Shafer (1976) extended and refined Dempster's work with respect to a belief function based on plausibility, thus improving the measure of subjective uncertainty. A further extension called evidential reasoning deals with information that is expected to be uncertain, imprecise and occasionally inaccurate (Lawrence, 1982). The Dempster-Shafer approach basically deviates from the Bayesian approach, which can be written as below: (Chang, 1991)

$$P(H_i | E_j) = P(E_j | H_i) P(H_i) / \sum_i P(E_j | H_i) P(H_i)$$

Where  $H_i$  is the set of possible outcome states (e.g., residential or commercial use) and  $E_j$  is the set of all possible condition states (e.g., high or low soil productivity). The prior probability  $P(H_i)$  and the conditional probability  $P(E_j | H_i)$  must be obtained before computing the posterior probability  $P(H_i | E_j)$ . There are two assumptions inherent in the Bayesian procedure. First, the elements of set E must be statistically independent. Second, the set E must be comprehensive in explaining the set of outcomes H. The Bayesian theory is valid only if these two assumptions are satisfied. Actually, the Dempster-Shafer theory can be used to avoid some limitations of Bayesian method on two points.

First, the Dempster-Shafer model overcomes the limitation of single evidence or a single hypothesis inherent in the Bayesian model. Dempster stated that basic probability assignments attach probabilities, not to individual possible states of the world, but more ambiguously to subsets of these states. When one seeks to compute an uncertain measure for a particular outcome, one is led in general to a pair of numbers: a 'belief' (or lower probability or minimal amount of probability assigned to the outcome) and a 'plausibility' (or upper probability or maximal amount of probability assigned to the outcome) (Dempster et al., 1988).

Second, the Dempster-Shafer approach improves the main deficiency of Bayesian probability, the treatment of ignorance. Probability theory must distribute an equal amount of probability even in ignorance. For example, if you have no prior knowledge, then you have to assume the probability of each possibility to be  $P = 1/N$ , where  $N$  is the total number of possibilities. This assignment of  $P$  is made by using the principle of indifference. The principle of indifference can lead to ridiculous results if applied without thinking. One example cited by Giarratano and Riley (1994) is either there is or is not oil under your house. By the above principle, if you have absolutely no other knowledge, there is a fifty percent probability of having oil under your house. Would you immediately expend all your savings to hire a drilling rig and start drilling in the kitchen?

The Dempster-Shafer theory assumes that there is a fixed set of mutually exclusive and exhaustive elements called the environment and represented by  $\Theta$ .

$$\Theta = \{ \theta_1, \theta_2, \dots, \theta_N \}$$

The Dempster-Shafer theory does not force belief to be assigned to ignorance or to refutation of a hypothesis. Any belief that is not assigned to a specific subset is considered no belief or nonbelief and just associated with the environment  $\Theta$ . Belief that refutes a hypothesis is disbelief, which is not nonbelief (Giarratano and Riley, 1994). According to Dempster-Shafer theory, the degree of belief assigned to support each hypothesis in a sample space that may sum up to a number less than one (Gordon et al., 1985). This new concept may make the Dempster-Shafer theory a better qualitative description of human judgment.

Other possible advantages of the Dempster-Shafer theory will be described below. Because measures of uncertainty are based on a belief function, human experts don't have to worry about prior probability and hence they can concentrate on the analysis of evidence when estimating the probability of each possibility. In contrast to a probability-based approach, Dempster-Shafer theory represents uncertainty in terms of one's degree of belief in the rules that describe some process or phenomenon, and not explicitly in the likelihood of some categorical outcome resulting from that process or phenomenon.

In order to apply the Dempster-Shafer approach to spatial land use suitability indexing model, some important concepts and terminology must be introduced. An environment is called a frame of discernment when its elements may be interpreted as possible answers and only one answer is correct. The term “discern” implies that it is possible to distinguish the one correct answer from all the other possible answers to a question. If an answer is not in the frame, then the frame must be enlarged to accommodate the additional element(s)  $\theta_{N+1}$ ,  $\theta_{N+2}$  and so forth. The frame of discernment (denoted  $\Theta$ ) is analogous to the sample space in probability theory. A frame of discernment can be viewed as a set or collection of different outcomes (the elements of that set). A focal element 'A' is a subset of frame  $\Theta$ . 'A' can be a singleton subset or contain multiple sub-elements. For example,  $\Theta = \{A, B\}$ ,  $A = \{a_1, a_2\}$ ,  $B = \{b_1\}$ . The basic probability value  $m(A)$  is understood to be the measure of belief (belief function) for subset A of  $\Theta$ . A measure of belief assigned to one subset of a focal element is also assigned to every other subset of that element. In the above example, if  $m(A) = 0.4$ , then  $m(a_1) = 0.4$  and  $m(a_2) = 0.4$ . In Dempster-Shafer theory, the basic probability assignment  $m: P(\Theta) \rightarrow [0,1]$  holds whenever :

$$m(\emptyset) = 0 \quad (B.1)$$

$$\sum_{A \subset \Theta} m(A) = 1 \quad (B.2)$$

The value assigned to  $m(\Theta)$  is the strength of the implication that the evidence does not mean anything. This means neither belief nor disbelief. The basic probability assignment is analogous to the testimony of an expert who gives reliable testimony  $1 - m(\Theta)$  and unreliable testimony  $m(\Theta)$ , which is also called the 'discount' on a expert's credibility (Cohen, 1986). The 'discount' allows one to quantify uncertainty about a judgment without specifically assigning probability to each possible outcome.

The value of  $m(Q)$  is the measure of belief that one assigns to Q, but it is not the total belief (belief function) assigned to Q. The belief function, Bel, is the total belief of a set and all its subsets. To obtain the measure of the total belief committed to Q, one must sum all the basic probability values  $m(R)$ , where R represents a subset of Q:

$$\text{Bel}(Q) = \sum_{R \subset Q} m(R) \quad (\text{B.3})$$

The belief function is also called a ‘lower probability’ by Dempster (1967). The belief function does not fully describe one's belief about the focal element  $Q$ , because  $\text{Bel}(Q)$  does not convey all the information about one's suspicion of  $Q$  (i.e., to what extent one believes its negation  $\bar{Q}$ ). Here Shafer introduced another measure called ‘degree of doubt’, the degree to which  $A$  is disbelieved or refuted.

$$\text{Dou}(Q) = \text{Bel}(\bar{Q}) \quad (\text{B.4})$$

In turn, the extent of doubt allows the specification of a measure of ‘plausibility’, the degree to which one fails to doubt  $Q$ , or the extent to which one finds  $Q$  credible or plausible (Shafer, 1976).

$$\text{Pl}(Q) = 1 - \text{Dou}(Q) \quad (\text{B.5})$$

In terms of the basic probability number, the plausibility  $\text{Pl}: P(\Theta) \rightarrow [0,1]$  can be defined as the following :

$$\begin{aligned} \text{Pl}(Q) &= 1 - \text{Bel}(\bar{Q}) = 1 - \sum_{R \subset \bar{Q}} m(R) \\ &= \sum_{R \cap Q \neq \emptyset} m(R) \end{aligned} \quad (\text{B.6})$$

From (B.3) to (B.6), one can conclude that  $\text{Pl}(Q) \geq \text{Bel}(Q)$

In evidential reasoning, the evidence will induce an evidential interval. The lower bound is called Bel in Dempster-Shafer theory. The upper bound is called the plausibility (Pl). The Bel is the minimum belief based on the evidences and the plausibility is the maximum belief we are willing to give. The ranges of Bel and Pl are  $0 \leq \text{Bel} \leq \text{Pl} \leq 1$ . Table B.1 (Wesley, 1986) lists some common evidential intervals, that is, [lower bound, upper bound].

Dempster's rule is most intuitive when it is expressed in terms of the basic probability numbers, especially they are depicted geometrically. Suppose  $m_i$  is the basic probability assignment for a belief function  $\text{Bel}_i$  over a frame  $\Theta$ , and the  $\text{Bel}_i$ 's focal elements are denoted by  $A_1, \dots, A_i, \dots, A_k$ . Then the probability values measured by the basic probability numbers  $m_i(A_1), \dots, m_i(A_i), \dots, m_i(A_k)$  can be depicted as segments of a line of total



length one, as in Figure B.1. A second belief function  $Bel_2$  can be depicted in the same way, with basic probability assignment  $m_2$  and focal elements  $B_1, \dots, B_j, \dots, B_l$  represented in Figure B.2. And Figure B.3 shows that how the two lines  $m_1$  and  $m_2$  can be orthogonally combined to obtain a square.

**Table B.1 Some common evidential intervals**

Evidential Interval	Meaning
[1, 1]	Completely true
[0, 0]	Completely false
[0, 1]	Completely ignorant
[Bel, 1]	Tends to support
[0, Pl]	Tends to refute
[Bel, Pl]	Tends to both support and refute

In order to combine  $Bel_1$  and  $Bel_2$ , the square in Figure B.3 can be viewed as the total probability measure. Suppose that  $Bel_1$  assigns vertical strips to its focal elements, while  $Bel_2$  assigns horizontal strips to its focal elements. For example, a vertical strip of measure  $m_1(A_i)$  is assigned to  $A_i$  by  $m_1$  and a horizontal strip of measure  $m_2(B_j)$  is assigned to  $B_j$  by  $m_2$ . The intersection of these two strips is measured by  $m_1(A_i)m_2(B_j)$ , and because it is committed both to  $A_i$  and  $B_j$ , the joint effect of  $Bel_1$  and  $Bel_2$  is assigned exactly to  $A_i \cap B_j$  (Shafer, 1976).

Similarly, the exact assignment of every rectangle in the Figure B.3 can be specified. A given subset  $S$  of  $\Theta$  may have more than one of these rectangles exactly assigned to it, in which case the total probability values exactly assigned to  $S$  will have measure

$$\sum_{A_i \cap B_j = S} m_1(A_i)m_2(B_j)$$

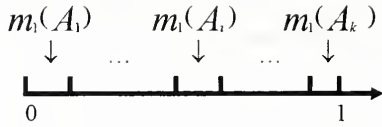


Figure B.1 The probability assignment based on A evidence

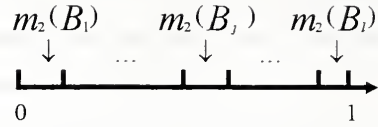


Figure B.2 The probability assignment based on B evidence

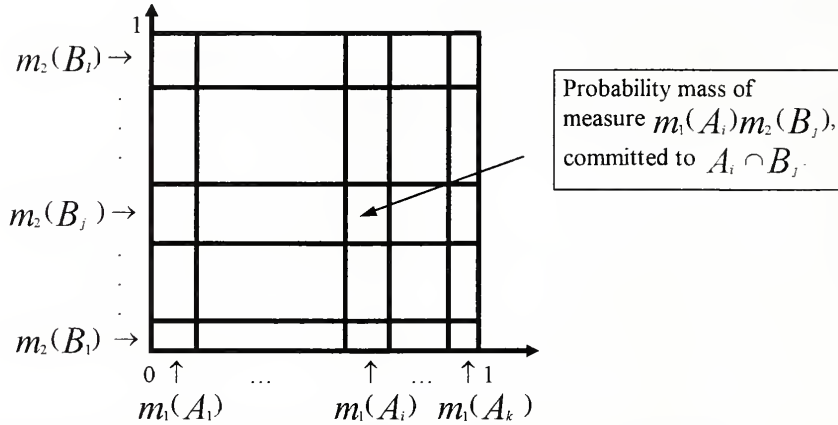


Figure B.3 The probability assignment based on A and B evidences

The only difficulty is that the above scheme may assign some rectangles of square to the empty set  $\emptyset$ . Hence, there may be a focal element  $A_i$  of  $Bel_1$  and a focal element  $B_j$  of  $Bel_2$  such that  $A_i \cap B_j = \emptyset$ , where

$$k = \sum_{A_i \cap B_j = \emptyset} m_1(A_i)m_2(B_j) \geq 0 \quad (\text{B.7})$$

where  $k$  indicates the amount of evidential conflict. The only remedy is to "discard" all the rectangles that are assigned to  $\emptyset$ . If not all the rectangles are discarded, the measures of the remaining rectangles can be inflated by multiplying them by the renormalizing factor  $K$ , where

$$K = (1 - \sum_{A_i \cap B_j = \emptyset} m_1(A_i)m_2(B_j))^{-1} = (1 - k)^{-1} \quad (\text{B.8})$$

so that the total probability value will again have measure one. This is called the normalization process. Because  $K$  increases with  $k$ , it can serve as a measure of the extent of the conflict. Shafer (1976) defined the measure of the conflict between  $Bel_1$  and  $Bel_2$  as the quantity

$$\log K = \log (1 / (1 - k)) = -\log (1 - k)$$

which he called the weight of conflict between  $Bel_1$  and  $Bel_2$ , and denoted it as  $\text{Con}(Bel_1, Bel_2)$ . If  $Bel_1$  and  $Bel_2$  do not conflict each other, meaning complete compatibility, then  $k = 0$  and  $\text{Con}(Bel_1, Bel_2) = 0$ . If  $Bel_1$  and  $Bel_2$  completely contradict each other, so that  $Bel_1 \oplus Bel_2$  does not exist, then  $k = 1$  and  $\text{Con}(Bel_1, Bel_2) = \infty$ . The operator ' $\oplus$ ' denotes Dempster's rule of combination. The belief function given by  $m$  is called the orthogonal sum of  $Bel_1$  and  $Bel_2$  and is denoted by  $Bel_1 \oplus Bel_2$ . If  $\sum_{A_i \cap B_j = \emptyset} m_1(A_i)m_2(B_j) < 1$  does not hold, then the orthogonal sum  $Bel_1 \oplus Bel_2$  does not exist.

Given several belief functions over the same frame of discernment, but based on distinct pieces of evidence, Dempster's rule of combination enables us to compute their orthogonal sum: a new belief function based on the combined evidence. Suppose  $\sum_{A \cap B_j = \emptyset} m_\alpha(A)m_\beta(B) < 1$ , then a new basic probability number of a proposition  $S$  from two independent sets of evidence  $\alpha$  and  $\beta$  is derived by the following equations:

$$m_{\alpha\beta}(S) = K \times \sum_{A \cap B = S} m_\alpha(A)m_\beta(B) \quad (\text{B.9})$$

$$K^{-1} = \sum_{A \cap B \neq \emptyset} m_\alpha(A)m_\beta(B) \quad (\text{B.10})$$

where  $\emptyset$  is the empty set, and  $A$  and  $B$  are possible subsets of the frame of discernment.  $m_\alpha(A)$  is the basic probability number that evidence  $\alpha$  implies hypothesis  $A$ , and similarly for  $m_\beta(B)$ .  $K$  is a factor used to renormalize the measure of the orthogonal sum.

Dempster's rule of combination ' $\oplus$ ' is used repeatedly until all evidence has been considered. For the general case having  $n$  pieces of evidence, denoted  $1, 2, \dots, n$

respectively, one can combine a collection of belief functions  $Bel_1, Bel_2, \dots, Bel_n$  over focal element of  $\Theta$  and compute the pairwise orthogonal sums.

$$Bel_{12} = Bel_1 \oplus Bel_2$$

$$Bel_{123} = Bel_1 \oplus Bel_2 \oplus Bel_3 = (Bel_1 \oplus Bel_2) \oplus Bel_3$$

.....

$$Bel_{12..n} = Bel_1 \oplus Bel_2 \oplus \dots \oplus Bel_n$$

Each stage  $i$  ( $i = 1, 2, \dots, n$ ) of Dempster rule's process should correspond to the addition of evidence supporting the  $Bel_i$  entered at that stage, and the final belief function  $Bel$  should represent the pooled evidence from the  $Bel_i$ . The body of evidence obtained by pooling several bodies of evidence will not depend on the order in which the pooling is done. Hence the equation of Dempster's rule with the pooling of evidence will be tenable only if the belief function  $Bel$  resulting from the pairwise combination of the  $Bel_i$  remains the same regardless of the order in which the  $Bel_i$  are entered.

One shortcoming in the Dempster-Shafer theory occurs with the application of normalization, which may lead to counterintuitive conclusions (Zadeh, 1984). The problem with normalization is that it ignores the belief that the object being considered does not exist. An example mentioned by Zadeh is the belief by two doctors, A and B, in a patient's illness. The diagnosis (or beliefs) in the patient's problem are as follows.

$$m_A(\text{meningitis}) = 0.99$$

$$m_A(\text{brain tumor}) = 0.01$$

$$m_B(\text{concussion}) = 0.99$$

$$m_B(\text{brain tumor}) = 0.01$$

Both doctors agree there is a very low probability, 0.01, of a brain tumor but greatly disagree on the more likely problem. The Dempster rule in this situation leads to a conclusion that the combined belief that this patient has brain tumor is 1. This result is very counterintuitive, because both doctors agree that the brain tumor is highly unlikely.

### Dempster-Shafer Evidence Theory for Land Suitability Analysis

Now, one example will demonstrate the application of the Dempster-Shafer evidence theory to land-use potential (or suitability) analysis. A frame of discernment (denoted  $\Theta$ ) contains eight factors (or evidences):

$$\Theta = \{ \theta_1, \theta_2, \dots, \theta_8 \}$$

= { Soil Productivity, Soil Limitation, Tendency to Flood, Forestation,  
Sanitary Sewer, Highway Accessibility, Railroads & Airports,  
Current & Expected Land Use }

By means of Dempster-Shafer equations (B.1) to (B.10), the probability measure of land use potential based on these eight factors can be calculated.

Example: Cell 21 has the following attributes for the eight factors.

1. 'low' Soil Productivity,
2. 'slight' Soil Limitation,
3. 'all other soils' in terms of Tendency to Flood,
4. 'no' Forested land,
5. '< 1000 ft.' from a Sanitary Sewer,
6. 'paved < 0.5 mile' for Accessibility,
7. '> 0.5 mile' from Railroads & Airports,
8. 'residential' as Current & Expected Land Use.

1	2	3	4	5	6
7	8	9	10	11	12
13	14	15	16	17	18
19	20	21	22	23	24
25	26	27	28	29	30
31	32	33	34	35	36

The rectangle shown above is 1 square mile in size, divided into 36 cells of equal area. What are the probability measures in cell 21 for 'Residential', 'Agricultural', 'Industrial', 'Commercial', and 'Open Space' use, respectively?

The eight factors and their probability judgments are represented in Table B.2. R, A, I, C, and O represent 'Residential', 'Agricultural', 'Industrial', 'Commercial', and 'Open Space' land uses, respectively.

Because the equation of Dempster-Shafer rule with the pooling of evidence remains the same regardless of the order in which the  $Bel_i$  are entered, a process using Dempster-Shafer rule is described as the following steps.

First of all, two factors (or evidences) are taken into account, namely 'low' Soil Productivity and 'slight' Soil Limitation. The beginning step is to calculate the  $K$  factor (see equation (B.10)) which is used to renormalize the inferred attributes.

$$\begin{aligned} K_{12}^{-1} &= 1 - \{m_1(R,I,C) * m_2(A,O) + m_1(O) * m_2(R,I,C)\} \\ &= 1 - (0.85 * 0.1 + 0.1 * 0.85) = 0.83 \end{aligned}$$

Then the inferred attributes can be obtained by using equation (B.9).

$$m_{12}(R) = 0$$

$$m_{12}(A) = 0$$

$$m_{12}(I) = 0$$

$$m_{12}(C) = 0$$

$$\begin{aligned} m_{12}(O) &= \{m_1(O) * m_2(A,O) + m_1(O) * m_2(\Theta)\} / K_{12}^{-1} \\ &= (0.1 * 0.1 + 0.1 * 0.05) / 0.83 = 0.01807 \end{aligned}$$

$$\begin{aligned} m_{12}(A,O) &= \{m_1(\Theta) * m_2(A,O)\} / K_{12}^{-1} \\ &= 0.05 * 0.1 / 0.83 = 0.006024 \end{aligned}$$

$$\begin{aligned} m_{12}(R,I,C) &= \{m_1(R,I,C) * m_2(R,I,C) + m_1(R,I,C) * m_2(\Theta) + m_1(\Theta) * m_2(R,I,C)\} / K_{12}^{-1} \\ &= (0.85 * 0.85 + 0.85 * 0.05 + 0.85 * 0.05) / 0.83 = 0.97289 \end{aligned}$$

$$\begin{aligned} m_{12}(\Theta) &= \{m_1(\Theta) * m_2(\Theta)\} / K_{12}^{-1} \\ &= 0.05 * 0.05 / 0.83 = 0.00301 \end{aligned}$$

**Table B.2 Probability measures for assessing land suitability (Dempster-Shafer Theory)**

1. Soil Productivity	Basic Probability Measure
low	$m_1(R,I,C) = 0.85, m_1(O) = 0.10, m_1(\Theta) = 0.05$
moderate	$m_1(R,I,C) = 0.50, m_1(A,O) = 0.30, m_1(\Theta) = 0.20$
high	$m_1(A) = 0.85, m_1(R,I,C,O) = 0.10, m_1(\Theta) = 0.05$
very high	$m_1(A) = 0.90, m_1(O) = 0.05, m_1(\Theta) = 0.05$
2. Soil Limitation	Basic Probability Measure
slight	$m_2(R,I,C) = 0.85, m_2(A,O) = 0.10, m_2(\Theta) = 0.05$
moderate	$m_2(R,I,C) = 0.48, m_2(A,O) = 0.48, m_2(\Theta) = 0.04$
severe	$m_2(R,I,C) = 0.10, m_2(A,O) = 0.85, m_2(\Theta) = 0.05$
3. Tendency to Flood	Basic Probability Measure
flooding soils	$m_3(A,O) = 0.95, m_3(\Theta) = 0.05$
all other soils	$m_3(\Theta) = 1.00$
4. Forested	Basic Probability Measure
yes	$m_4(O) = 0.55, m_4(R) = 0.35, m_4(A,C) = 0.05, m_4(\Theta) = 0.05$
no	$m_4(\Theta) = 1.00$
5. Sanitary Sewer	Basic Probability Measure
< 1000 ft.	$m_5(R,I,C) = 0.95, m_5(\Theta) = 0.05$
1000 ft - 1.5 mile	$m_5(I,C) = 0.70, m_5(O) = 0.10, m_5(\Theta) = 0.20$
> 1.5 mile	$m_5(A) = 0.75, m_5(O) = 0.20, m_5(\Theta) = 0.05$
6. Accessibility	Basic Probability Measure
paved < 0.5 mile	$m_6(R) = 0.50, m_6(I,C) = 0.45, m_6(\Theta) = 0.05$
paved > 0.5 mile	$m_6(R) = 0.45, m_6(A,O) = 0.40, m_6(\Theta) = 0.15$
major intersection or State/Federal R.O.W.	$m_6(I,C) = 0.85, m_6(A,O) = 0.10, m_6(\Theta) = 0.05$
Interstate R.O.W	$m_6(I,C) = 0.30, m_6(A,O) = 0.10, m_6(\Theta) = 0.60$



**Table B.2 (continued)**

7.Railroads & Airports	Basic Probability Measure
< 300 ft & proximate	$m_7(I) = 0.80, m_7(O) = 0.15, m_7(\Theta) = 0.05$
300 ft - 0.5 mile	$m_7(I) = 0.70, m_7(\Theta) = 0.30$
> 0.5 mile	$m_7(\Theta) = 1.00$
8.Current & Expected Land Use	Basic Probability Measure
residential	$m_8(R) = 0.99, m_8(\Theta) = 0.01$
agricultural	$m_8(A) = 0.85, m_8(\Theta) = 0.15$
industrial	$m_8(I) = 0.99, m_8(\Theta) = 0.01$
commercial	$m_8(C) = 0.99, m_8(\Theta) = 0.01$
open space	$m_8(O) = 0.85, m_8(\Theta) = 0.15$

Next, the evidence ‘all other soils’ for Tendency to Flood will be included by applying the same equations.

$$K_{123}^{-1} = 1 - 0 = 0$$

$$m_{123}(R) = 0$$

$$m_{123}(A) = 0$$

$$m_{123}(I) = 0$$

$$m_{123}(C) = 0$$

$$m_{123}(O) = 0.01807 * 1.0/1 = 0.01807$$

$$m_{123}(A,O) = 0.00602 * 1.0/1 = 0.00602$$

$$m_{123}(R,I,C) = 0.97289 * 1.0/1 = 0.97289$$

$$m_{123}(\Theta) = 0.00301 * 1.0/1 = 0.00301$$

Then include the evidence ‘no’ for Forested land and apply the same equations.

$$K_{1234}^{-1} = 1 - 0 = 1$$

$$m_{1234}(R) = 0$$

$$m_{1234}(A) = 0$$

$$m_{1234}(I) = 0$$

$$m_{1234}(C) = 0$$

$$m_{1234}(O) = 0.01807 * 1.0/1 = 0.01807$$

$$m_{1234}(A,O) = 0.00602 * 1.0/1 = 0.00602$$

$$m_{1234}(R,I,C) = 0.97289 * 1.0/1 = 0.97289$$

$$m_{1234}(\Theta) = 0.00301 * 1.0/1 = 0.00301$$

Next, the evidence '< 1000 ft' from Sanitary & Sewer is included.

$$K_{12345}^{-1} = 1 - (0.01807 * 0.95 + 0.00602 * 0.95) = 0.97711$$

$$m_{12345}(R) = 0$$

$$m_{12345}(A) = 0$$

$$m_{12345}(I) = 0$$

$$m_{12345}(C) = 0$$

$$m_{12345}(O) = 0.01807 * 0.05/0.97711 = 0.000925$$

$$m_{12345}(A,O) = 0.00602 * 0.05/0.97711 = 0.000308$$

$$m_{12345}(R,I,C) = (0.97289 * 0.95 + 0.97289 * 0.05 + 0.00301 * 0.05)/0.97711 = 0.995835$$

$$m_{12345}(\Theta) = 0.00301 * 0.05/0.97711 = 0.000154$$

Next, include the evidence 'paved < 0.5 mile' for Accessibility,

$$K_{123456}^{-1} = 1 - [(0.000925 + 0.000308) * 0.5 + (0.000925 + 0.000308) * 0.45] = 0.998828$$

$$m_{123456}(R) = (0.995835 + 0.000154) * 0.5/0.998828 = 0.498579$$

$$m_{123456}(A) = 0$$

$$m_{123456}(I) = 0$$

$$m_{123456}(C) = 0$$

$$m_{123456}(O) = 0.000925 * 0.05/0.998828 = 0.000046$$

$$m_{123456}(A,O) = 0.000308 * 0.05/0.998828 = 0.000015$$

$$m_{123456}(I,C) = (0.995835 + 0.000154) * 0.45/0.998828 = 0.448721$$

$$m_{123456}(R,I,C) = 0.995835 * 0.05/0.998828 = 0.04985$$

$$m_{123456}(\Theta) = 0.000154 * 0.05/0.998828 = 0.0000077$$

Then, include the evidence '> 0.5 mile' from Railroads & Airports.

$$K_{1234567}^{-1} = 1 - 0 = 1$$

$$m_{1234567}(R) = 0.498578 * 1.0/1 = 0.498579$$

$$m_{1234567}(A) = 0$$

$$\begin{aligned}
m_{1234567}(I) &= 0 \\
m_{1234567}(C) &= 0 \\
m_{1234567}(O) &= 0.000046 * 1.0/1 = 0.000046 \\
m_{1234567}(A,O) &= 0.000015 * 1.0/1 = 0.000015 \\
m_{1234567}(I,C) &= 0.44872 * 1.0/1 = 0.448721 \\
m_{1234567}(R,I,C) &= 0.04985 * 1.0/1 = 0.04985 \\
m_{1234567}(\Theta) &= 0.0000077 * 1.0/1 = 0.0000077
\end{aligned}$$

Finally, the evidence ‘residential’ for Current & Expected Land Use will be included.

$$\begin{aligned}
K_{12345678}^{-1} &= 1 - [(0.000046 + 0.000015 + 0.448721) * 0.99] = 0.555706 \\
m_{12345678}(R) &= [0.498579 * (0.99 + 0.01) + 0.04985 * 0.99] / 0.555706 = 0.986008 \\
m_{12345678}(A) &= 0 \\
m_{12345678}(I) &= 0 \\
m_{12345678}(C) &= 0 \\
m_{12345678}(O) &= 0.000046 * 0.01 / 0.555706 = 0.0000008 \\
m_{12345678}(A,O) &= 0.000015 * 0.01 / 0.555706 = 0.0000003 \\
m_{12345678}(I,C) &= 0.448721 * 0.01 / 0.555706 = 0.008075 \\
m_{12345678}(R,I,C) &= 0.04985 * 0.01 / 0.555706 = 0.000897 \\
m_{12345678}(\Theta) &= 0.0000077 * 0.01 / 0.555706 = 0.0000001
\end{aligned}$$

These nine final basic probability measures based on Dempster's rule of combination are then used to calculate the belief function and plausibility (refer to equations B.3 and B.6) for each inferred attribute. The final results in Table B.3 and B.4 show the “Residential” use having the highest probability measure among all different land uses. This indicates that the greatest potential use of cell 21 is “residential” use, meaning that the current land use of cell 21 is consistent with its best potential use.

The best potential use of each study area derived from the Dempster's rule can be viewed as the most suitable use for this area. Furthermore, the Dempster-Shafer process can point out some inconsistencies between the current land use and its potential land use for each study area. Using the above example, let us keep all factors having the same attributes, except the attribute for “Current & Expected Land Use”.

**Table B.3 Calculation of lower probability**

Land Use Category	Belief (Lower Probability)
R	$m(R) = 0.986008$
A	$m(A) = 0$
I	$m(I) = 0$
C	$m(C) = 0$
O	$m(O) = 0.0000008$
A,O	$m(A)+m(O)+m(A,O) = 0.0000011$
I,C	$m(I) + m(C)+m(I,C) = 0.008075$
R,I,C	$m(R)+m(I)+m(C)+m(I,C)+m(R,I,C) = 0.99498$

**Table B.4 Calculation of upper probability**

Land Use Category	Plausibility (Upper Probability)
R	$m(R) = 0.9869050$
A	$m(A) = 0.0000004$
I	$m(I) = 0.0089721$
C	$m(C) = 0.0089721$
O	$m(O) = 0.0000012$
A,O	$m(A)+m(O)+m(A,O)+m(\Theta) = 0.0000012$
I,C	$m(I) + m(C)+m(I,C)+m(R,I,C)+m(\Theta) = 0.0089721$
R,I,C	$m(R)+m(I)+m(C)+m(I,C)+m(R,I,C)+m(\Theta) = 0.99498$

Assume that the current use of cell 21 could be ‘Agricultural’ or ‘Industrial’ or ‘Commercial’ or ‘Open space’, not ‘Residential’. The results of all these four cases are described below.

1) If the attribute for Current & Expected Land Use for cell 21 is ‘Agricultural’, then

$$K_{12345678}^{-1} = 1 - [(0.498579+0.000046+0.448721+0.04985)*0.85] = 0.152383$$

$$m_{12345678}(R) = 0.498579*0.15/0.152383 = 0.49078$$

$$m_{12345678}(A) = (0.000015+0.0000077)*0.85/0.152383 = 0.000088$$

$$m_{12345678}(I) = 0$$

$$m_{12345678}(C) = 0$$

$$m_{12345678}(O) = 0.000046*0.15/0.152383 = 0.000045$$

$$m_{12345678}(A,O) = 0.000015*0.15/0.152383 = 0.000015$$

$$m_{12345678}(I,C) = 0.448721*0.15/0.152383 = 0.441703$$

$$m_{12345678}(R,I,C) = 0.04985*0.15/0.152383 = 0.04907$$

$$m_{12345678}(\Theta) = 0.00000077 * 0.15 / 0.152383 = 0.00000076$$

2) If the evidence 'Industrial' for Current & Expected Use will be included, then

$$K_{12345678}^{-1} = 1 - [(0.498579 + 0.000046 + 0.000015) * 0.99] = 0.506346$$

$$m_{12345678}(R) = 0.498579 * 0.01 / 0.506346 = 0.009847$$

$$m_{12345678}(A) = 0$$

$$m_{12345678}(I) = (0 + 0.448721 + 0.04985 + 0.00000077) * 0.99 / 0.506346 = 0.974799$$

$$m_{12345678}(C) = 0$$

$$m_{12345678}(O) = 0.000046 * 0.01 / 0.506346 = 0.0000009$$

$$m_{12345678}(A, O) = 0.000015 * 0.01 / 0.506346 = 0.0000003$$

$$m_{12345678}(I, C) = 0.448721 * 0.01 / 0.506346 = 0.008862$$

$$m_{12345678}(R, I, C) = 0.04985 * 0.01 / 0.506346 = 0.000984$$

$$m_{12345678}(\Theta) = 0.00000077 * 0.01 / 0.506346 = 0.000000016$$

3) If cell 21's Current & Expected Use is 'Commercial', then

$$K_{12345678}^{-1} = 1 - [(0.498579 + 0.000046 + 0.000015) * 0.99] = 0.506346$$

$$m_{12345678}(R) = 0.498579 * 0.01 / 0.506346 = 0.009847$$

$$m_{12345678}(A) = 0$$

$$m_{12345678}(I) = 0$$

$$m_{12345678}(C) = (0 + 0.448721 + 0.04985 + 0.00000077) * 0.99 / 0.506346 = 0.974799$$

$$m_{12345678}(O) = 0.000046 * 0.01 / 0.506346 = 0.0000009$$

$$m_{12345678}(A, O) = 0.000015 * 0.01 / 0.506346 = 0.0000003$$

$$m_{12345678}(I, C) = 0.448721 * 0.01 / 0.506346 = 0.008862$$

$$m_{12345678}(R, I, C) = 0.04985 * 0.01 / 0.506346 = 0.000984$$

$$m_{12345678}(\Theta) = 0.00000077 * 0.01 / 0.506346 = 0.000000016$$

4) If the evidence 'Open Space' for Current & Expected Use will be included, then

$$K_{12345678}^{-1} = 1 - [(0.498579 + 0.448721 + 0.04985) * 0.85] = 0.152422$$

$$m_{12345678}(R) = 0.498579 * 0.15 / 0.152422 = 0.490656$$

$$m_{12345678}(A) = 0$$

$$m_{12345678}(I) = 0$$

$$m_{12345678}(C) = 0$$

$$m_{12345678}(O) = [0.000046*(1)+(0.000015+0.00000077)*0.85]/0.152422 = 0.000389$$

$$m_{12345678}(A,O) = 0.000015*0.15/0.152422 = 0.000015$$

$$m_{12345678}(I,C) = 0.448721*0.15/0.152422 = 0.441591$$

$$m_{12345678}(R,I,C) = 0.04985*0.15/0.152422 = 0.049058$$

$$m_{12345678}(\Theta) = 0.00000077*0.15/0.152422 = 0.00000076$$

In cases 2 and 3, the results show agreement between the current use and potential use for cell 21. The results in case 1 and case 4 indicate inconsistencies between the potential land use and its current use. Using Dempster's rule of combination, acceptable potential uses for cell 21 are 'residential' or 'commercial' or 'industrial'. 'Agricultural' and 'open space' use, however, would be not suitable for cell 21, based on its eight factors.

Based on examples like these, the Dempster-Shafer approach seems capable of providing helpful directions to a land use planner, or owner or developer of land, helping to avoid undesirable uses of land.







